

Lineare Algebra und Geometrie

Stefan Haller

Wintersemester 2011/12

Sommersemester 2012

Wintersemester 2012/13

Dies ist ein Vorlesungsskriptum zu einer Lehrveranstaltung, die ich im Wintersemester 2011/12, Sommersemester 2012 und Wintersemester 2012/13 an der Universität Wien gehalten habe. Es steht unter

<http://www.mat.univie.ac.at/~stefan/LA2.html>

zur Verfügung.

Wien, am 26. Jänner 2013

INHALTSVERZEICHNIS

I. Einleitung	3
I.1. Ein lineares Gleichungssystem	3
I.2. Geometrische Interpretation	10
II. Vektorräume und lineare Abbildungen	13
II.1. Vektorräume	13
II.2. Teilräume	19
II.3. Lineare Abbildungen	22
II.4. Matrizen	29
II.5. Summen und Komplemente	36
II.6. Quotientenräume	42
III. Basen	45
III.1. Erzeugendensysteme	45
III.2. Lineare Unabhängigkeit	49
III.3. Basen	53
III.4. Dualräume	61
IV. Endlich-dimensionale Vektorräume	69
IV.1. Dimension	69
IV.2. Dimensionsformeln	74
IV.3. Rang von Matrizen	81

IV.4. Inhomogene Gleichungssysteme	97
IV.5. Matrizeninversion	100
IV.6. Basisdarstellung	105
V. Determinanten	119
V.1. Determinantenfunktionen	119
V.2. Determinanten und Gleichungssysteme	130
V.3. Permutationen und Leibniz'sche Formel	132
V.4. Determinanten von Endomorphismen	135
V.5. Orientierung reeller Vektorräume	137
VI. Eigenwerte und Eigenvektoren	141
VI.1. Diagonalisierbarkeit	141
VI.2. Charakteristisches Polynom	154
VI.3. Jordan'sche Normalform	166
VI.4. Reelle Normalformen	185
VII. Euklidische und unitäre Vektorräume	193
VII.1. Symmetrische Bilinearformen	193
VII.2. Innere Produkte	213
VII.3. Normale und selbstadjungierte Operatoren	228
VII.4. Isometrien	236
VII.5. Positivität	241
VII.6. Moore–Penrose Pseudoinverse	250
VIII. Affine und projektive Geometrie	257
VIII.1. Affine Räume und Abbildungen	257
VIII.2. Affine Quadriken	268
VIII.3. Projektive Räume und Abbildungen	281
VIII.4. Projektive Quadriken	299
IX. Multilineare Algebra	303
IX.1. Produkte und direkte Summen	303
IX.2. Tensorprodukt	308
IX.3. Erweiterung der Skalare	315
IX.4. Tensoralgebra	317
IX.5. Grassmann Algebra	320
IX.6. Symmetrische Algebra	326
Literatur	331
Übungen zu “Einführung in die lineare Algebra und Geometrie”	1
Übungen zu “Lineare Algebra und Geometrie 1”	1
Übungen zu “Lineare Algebra und Geometrie 2”	1

I. Einleitung

Die lineare Algebra beschäftigt sich mit dem Studium linearer Abbildungen zwischen Vektorräumen und ist somit auch das geeignete Werkzeug um lineare Gleichungssysteme zu untersuchen. Lösbarkeitsfragen zu linearen Gleichungssystemen haben sehr einfache qualitative Antworten. Selbst große lineare Gleichungssysteme können effizient mit Computerunterstützung gelöst werden, diese Algorithmen haben mittlerweile eine Unzahl an konkreten Anwendungen gefunden. Lineare Gleichungssysteme sind aber auch von theoretischem Interesse. Etwa lassen sich interessantere, d.h. nicht-lineare Gleichungen oft erfolgreich durch lineare approximieren und die Lösbarkeit der linearen Approximation erlaubt manchmal Rückschlüsse auf das ursprüngliche, nicht-lineare Problem.

In dieser Einleitung wollen wir einige der Resultate der folgenden Kapitel anhand eines konkreten Beispiels illustrieren.

I.1. Ein lineares Gleichungssystem. Betrachte folgende Gleichungen:

$$\begin{array}{rcccccccc} x_1 & -x_2 & -9x_3 & +x_4 & +x_5 & +x_6 & +3x_7 & = & y_1 \\ 2x_1 & -x_2 & -9x_3 & +9x_4 & +3x_5 & +7x_6 & +11x_7 & = & y_2 \\ -x_1 & +3x_2 & +27x_3 & +13x_4 & +2x_5 & +9x_6 & +9x_7 & = & y_3 \\ x_1 & -2x_2 & -18x_3 & -6x_4 & +2x_5 & -4x_6 & +2x_7 & = & y_4 \\ 2x_1 & +x_2 & +9x_3 & +23x_4 & +8x_5 & +17x_6 & +27x_7 & = & y_5 \end{array} \quad (\text{I.1})$$

Für welche reelle Zahlen y_1, y_2, y_3, y_4 und y_5 besitzt dieses Gleichungssystem eine Lösung, d.h. existieren reelle Zahlen $x_1, x_2, x_3, x_4, x_5, x_6$ und x_7 , die allen fünf Gleichungen genügen? Wieviele solche Lösungen gibt es, und wie lassen sie sich finden?

Wir wollen den rechnerischen Aspekt auf später verschieben und zunächst erläutern, was sich mit den Methoden der linearen Algebra über dieses Gleichungssystem sagen lässt, ohne überhaupt zu rechnen zu beginnen. Um die Notation zu vereinfachen definieren wir mit Hilfe der linken Seite von (I.1) eine Abbildung $\psi: \mathbb{R}^7 \rightarrow \mathbb{R}^5$,

$$\psi \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \\ x_7 \end{pmatrix} := \begin{pmatrix} x_1 - x_2 - 9x_3 + x_4 + x_5 + x_6 + 3x_7 \\ 2x_1 - x_2 - 9x_3 + 9x_4 + 3x_5 + 7x_6 + 11x_7 \\ -x_1 + 3x_2 + 27x_3 + 13x_4 + 2x_5 + 9x_6 + 9x_7 \\ x_1 - 2x_2 - 18x_3 - 6x_4 + 2x_5 - 4x_6 + 2x_7 \\ 2x_1 + x_2 + 9x_3 + 23x_4 + 8x_5 + 17x_6 + 27x_7 \end{pmatrix}.$$

Das Gleichungssystem (I.1) lässt sich damit in kompakter Form schreiben,

$$\psi(x) = y \quad (\text{I.2})$$

wobei $y \in \mathbb{R}^5$ und $x \in \mathbb{R}^7$. Die Linearität der Gleichungen spiegelt sich in folgender Eigenschaft der Abbildung ψ wider,

$$\psi(x + \tilde{x}) = \psi(x) + \psi(\tilde{x}) \quad \text{und} \quad \psi(\lambda x) = \lambda\psi(x), \quad (\text{I.3})$$

für alle $x, \tilde{x} \in \mathbb{R}^7$ und $\lambda \in \mathbb{R}$. Solche Abbildungen werden wir später als *lineare Abbildungen* bezeichnen, vgl. Übungsaufgabe 1.

Wir betrachten zunächst den Fall $y = 0$, und stellen folgende Frage: Wie lässt sich die Lösungsmenge des Gleichungssystems $\psi(x) = 0$ beschreiben? Dieses Gleichungssystem, $\psi(x) = 0$, wird das assoziierte *homogene Gleichungssystem* genannt, wir bezeichnen seine Lösungsmenge mit

$$L := \{x \in \mathbb{R}^7 : \psi(x) = 0\}.$$

Die Menge L besteht daher aus allen Vektoren $x \in \mathbb{R}^7$, deren Komponenten x_1, \dots, x_7 folgendes Gleichungssystem erfüllen:

$$\begin{array}{ccccccc} x_1 & -x_2 & -9x_3 & +x_4 & +x_5 & +x_6 & +3x_7 & = & 0 \\ 2x_1 & -x_2 & -9x_3 & +9x_4 & +3x_5 & +7x_6 & +11x_7 & = & 0 \\ -x_1 & +3x_2 & +27x_3 & +13x_4 & +2x_5 & +9x_6 & +9x_7 & = & 0 \\ x_1 & -2x_2 & -18x_3 & -6x_4 & +2x_5 & -4x_6 & +2x_7 & = & 0 \\ 2x_1 & +x_2 & +9x_3 & +23x_4 & +8x_5 & +17x_6 & +27x_7 & = & 0 \end{array} \quad (\text{I.4})$$

Da offensichtlich $\psi(0) = 0$ gilt, besitzt das homogene Gleichungssystem wenigstens die *triviale Lösung* $x = 0$, die Lösungsmenge ist daher nicht leer, es gilt $0 \in L$. Darüber hinaus hat L folgende Eigenschaft: Sind x und \tilde{x} aus L , dann liegt auch ihre Summe $x + \tilde{x}$ in L , und für jede reelle Zahl λ gilt auch $\lambda x \in L$. Dies folgt sofort aus der Linearität von ψ , siehe (I.3) und Übungsaufgabe 2. Teilmengen mit dieser Eigenschaft werden wir später als *Teilräume* bezeichnen.

Nach den Resultaten in Kapitel IV besitzen Teilräume stets Basen, d.h. es existieren Vektoren $b_1, \dots, b_l \in L$, sodass sich jedes Element $x \in L$ in der Form

$$x = s_1 b_1 + \dots + s_l b_l$$

schreiben lässt mit eindeutig bestimmten Skalaren $s_1, \dots, s_l \in \mathbb{R}$. Der Lösungsraum L kann daher durch l reelle Zahlen parametrisieren. Genauer, ist die Abbildung

$$\phi: \mathbb{R}^l \xrightarrow{\cong} L, \quad \phi \begin{pmatrix} s_1 \\ \vdots \\ s_l \end{pmatrix} = s_1 b_1 + \dots + s_l b_l \quad (\text{I.5})$$

eine Bijektion. Manchmal wird dies auch als eine *Parameterdarstellung* von L bezeichnet. Die Surjektivität von ϕ bedeutet gerade, dass wir so alle Lösungen von (I.4) erhalten, d.h.

$$L = \{s_1 b_1 + \dots + s_l b_l \mid s_1, \dots, s_l \in \mathbb{R}\},$$

und die Injektivität besagt, dass die Parametrisierung ϕ jede Lösung nur einmal liefert. Beachte, dass auch die Parametrisierung ϕ linear ist, denn es gilt offensichtlich $\phi(s + \tilde{s}) = \phi(s) + \phi(\tilde{s})$ und $\phi(\lambda s) = \lambda\phi(s)$, für alle $s, \tilde{s} \in \mathbb{R}^l$ und $\lambda \in \mathbb{R}$,

vgl. Übungsaufgabe 1. Dies bedeutet, dass Addition und Skalarmultiplikation in L genau der üblichen Addition und Skalarmultiplikation in \mathbb{R}^l entsprechen, d.h. ϕ respektiert die lineare Struktur. Wir werden solche Abbildungen später als *lineare Isomorphismen* bezeichnen.

Im Allgemeinen wird L viele verschiedene Basen besitzen, nach den Ergebnissen in Kapitel IV ist die Anzahl der Basisvektoren jedoch immer dieselbe. Diese Zahl l wird als *Dimension* des Teilraums bezeichnet, wir schreiben

$$\dim(L) = l.$$

Da L ein Teilraum von \mathbb{R}^7 ist, muss jedenfalls $0 \leq l \leq 7$ gelten, wie wir später sehen werden. Das *Gauß'sche Eliminationsverfahren* liefert einen effizienten Algorithmus zur Bestimmung einer Basis b_1, \dots, b_l von L , wir werden dies unten an unserem Beispiel ausführen, vgl. (I.14).

Die Parameterdarstellung ist gut geeignet (alle) Lösungen des Gleichungssystems anzugeben. Soll umgekehrt überprüft werden ob ein gegebenes $x \in \mathbb{R}^7$ das Gleichungssystem (I.4) erfüllt, dann lässt sich dies durch direktes Einsetzen in die Gleichungen sofort beantworten. Dabei stellt sich allerdings die Frage, ob es wirklich notwendig ist alle Gleichungen zu überprüfen oder ob vielleicht weniger Gleichungen ausreichen. In unserem Beispiel (I.4) gilt etwa: Elf mal die erste Gleichung minus drei mal die zweite plus drei mal die dritte ergibt genau die fünfte Gleichung. Die letzte Gleichung ist daher überflüssig, sie ist eine Konsequenz der vorangehenden. Auch die vierte Gleichung lässt sich übrigens durch geschicktes Kombinieren aus den ersten drei Gleichungen herleiten, auch sie ist redundant. Es stellt sich nun heraus, dass es stets möglich ist L durch ein homogenes lineares Gleichungssystem mit $7 - l$ vielen Gleichungen zu beschreiben. Nebenbei, zumindest ein solches Gleichungssystem lässt sich stets durch Weglassen geeigneter Gleichungen in (I.4) erhalten. Wieviele und welche wegzulassen sind ist jedoch ohne Rechnung nicht klar. Wir werden unten ein möglichst einfaches Gleichungssystem für L angeben, siehe (I.13).

Nun da wir die Lösungen des homogenen Systems (I.4) zumindest qualitativ gut verstehen, stellen wir folgende naheliegende Frage: Für welche y besitzt das Gleichungssystem (I.1) wenigstens eine Lösung, bzw. wie lässt sich die Menge dieser y , d.h. das Bild der Abbildung ψ ,

$$W := \psi(\mathbb{R}^7) = \{y \in \mathbb{R}^5 \mid \exists x \in \mathbb{R}^7 : \psi(x) = y\},$$

gut beschreiben? Offensichtlich gilt $0 \in W$, denn $\psi(0) = 0$. Aus (I.3) folgt sofort, dass W einen Teilraum von \mathbb{R}^5 bildet, d.h. sind $y, \tilde{y} \in W$, dann gilt auch $y + \tilde{y} \in W$ und $\lambda y \in W$, für alle $\lambda \in \mathbb{R}$, vgl. Übungsaufgabe 2. Nach den oben erwähnten Resultaten besitzt W daher eine Basis, d.h. es existieren Vektoren $w_1, \dots, w_k \in W$, sodass sich jedes $y \in W$ auf eindeutige Weise in der Form

$$y = t_1 w_1 + \dots + t_k w_k$$

schreiben lässt, für geeignete Skalare $t_1, \dots, t_k \in \mathbb{R}$. Auch W lässt sich daher in linearer Weise parametrisieren,

$$\mathbb{R}^k \xrightarrow{\cong} W, \quad \begin{pmatrix} t_1 \\ \vdots \\ t_k \end{pmatrix} \mapsto t_1 w_1 + \dots + t_k w_k, \quad (\text{I.6})$$

ist eine Bijektion, d.h. ein linearer Isomorphismus. Insbesondere gilt

$$W = \{t_1 w_1 + \dots + t_k w_k \mid t_1, \dots, t_k \in \mathbb{R}\}.$$

Soll umgekehrt von einem gegebenen $y \in \mathbb{R}^5$ überprüft werden, ob es in W liegt, dann wäre ein Gleichungssystem für W besser geeignet. Analog zu L , lässt sich W durch ein homogenes Gleichungssystem mit $5 - k$ vielen Gleichungen in den Variablen y_1, \dots, y_5 beschreiben. Wir werden unten ein möglichst einfaches solches Gleichungssystem angeben, siehe (I.10), und auch eine Basis w_1, \dots, w_k von W explizit bestimmen, vgl. (I.11). Nebenbei, wenigstens eine der vielen Basen von W lässt sich gewinnen, indem wir die Koeffizienten des Gleichungssystems (I.4) als Vektoren in \mathbb{R}^5 auffassen, d.h. die Vektoren

$$\begin{pmatrix} 1 \\ 2 \\ -1 \\ 1 \\ 2 \end{pmatrix}, \quad \begin{pmatrix} -1 \\ -1 \\ 3 \\ -2 \\ 1 \end{pmatrix}, \quad \begin{pmatrix} -9 \\ -9 \\ 27 \\ -18 \\ 9 \end{pmatrix}, \quad \begin{pmatrix} 1 \\ 9 \\ 13 \\ -6 \\ 23 \end{pmatrix}, \quad \begin{pmatrix} 1 \\ 3 \\ 2 \\ 2 \\ 8 \end{pmatrix}, \quad \begin{pmatrix} 1 \\ 7 \\ 9 \\ -4 \\ 17 \end{pmatrix}, \quad \begin{pmatrix} 3 \\ 11 \\ 9 \\ 2 \\ 27 \end{pmatrix},$$

betrachten und geeignete davon auswählen. Wieviele und welche dies sein sollen ist jedoch nicht ganz offensichtlich.

Wir haben oben die Anzahl der Basisvektoren von W mit k bezeichnet, d.h.

$$\dim(W) = k.$$

Da W ein Teilraum von \mathbb{R}^5 ist, muss jedenfalls $0 \leq k \leq 5$ gelten. Interessanter Weise sind die Dimensionen von L und W nicht unabhängig voneinander. Nach einer *Dimensionsformel* in Kapitel IV muss stets

$$l + k = \dim(L) + \dim(W) = \dim(\mathbb{R}^7) = 7 \quad (\text{I.7})$$

gelten. Insbesondere folgt $l = 7 - k \geq 7 - 5 = 2$, d.h.

$$l = \dim(L) \geq 2.$$

Es sind daher mindestens zwei reelle Parameter notwendig, um den Lösungsraum L in linearer Weise zu parametrisieren.

Wir wenden uns nun dem ursprünglichen Gleichungssystem (I.1) zu. Zu gegebenem $y \in W$ fragen wir: Wie lässt sich die Lösungsmenge des Systems $\psi(x) = y$ gut beschreiben? Wir geben auch ihr einen Namen,

$$L_y := \{x \in \mathbb{R}^7 \mid \psi(x) = y\}.$$

Da $y \in W$ gibt es zumindest eine Lösung $\xi_y \in L_y$, d.h. $\psi(\xi_y) = y$. Aus der Linearität von ψ , siehe (I.3), folgt nun sofort, dass wir alle Lösungen des Systems $\psi(x) = y$ erhalten, in dem wir zu der *speziellen Lösung* ξ_y die Lösungen des homogenen Systems addieren, d.h.

$$L_y = \{\xi_y + x \mid \psi(x) = 0\} = \xi_y + L.$$

Die Lösungsmenge L_y entsteht daher aus L durch Verschieben um ξ_y . Somit lässt sich auch L_y parametrisieren,

$$\phi_y: \mathbb{R}^l \xrightarrow{\cong} L_y, \quad \phi_y \begin{pmatrix} s_1 \\ \vdots \\ s_l \end{pmatrix} = \xi_y + s_1 b_1 + \cdots + s_l b_l, \quad (\text{I.8})$$

ist eine Bijektion, siehe Übungsaufgabe 6.

Beachte, dass L_y im Allgemeinen keinen Teilraum bildet, denn jeder Teilraum muss den Nullvektor enthalten, aber i.A. wird $0 \notin L_y$ gelten. Auch die Parametrisierung ϕ_y ist i.A. nicht linear.¹ Allerdings ist sie *affin*, d.h. es gilt

$$\phi_y(\lambda s + (1 - \lambda)\tilde{s}) = \lambda\phi_y(s) + (1 - \lambda)\phi_y(\tilde{s}),$$

für alle $s, \tilde{s} \in \mathbb{R}^l$ und $\lambda \in \mathbb{R}$. Die Parametrisierung ϕ_y respektiert daher *Konvexkombinationen*.

Sind $y, \tilde{y} \in W$ und $\xi_y \in L_y$ sowie $\xi_{\tilde{y}} \in L_{\tilde{y}}$ spezielle Lösungen, d.h. $\psi(\xi_y) = y$ und $\psi(\xi_{\tilde{y}}) = \tilde{y}$, dann gilt auch $\psi(\xi_y + \xi_{\tilde{y}}) = y + \tilde{y}$ und $\psi(\lambda\xi_y) = \lambda y$, siehe (I.3), d.h. $\xi_y + \xi_{\tilde{y}}$ ist eine spezielle Lösung des Gleichungssystems $\psi(x) = y + \tilde{y}$ und $\lambda\xi_y$ ist eine spezielle Lösung von $\psi(x) = \lambda y$. Tatsächlich existiert stets eine lineare Abbildung $\xi: \mathbb{R}^5 \rightarrow \mathbb{R}^7$, sodass $\psi(\xi(y)) = y$, für alle $y \in W$. In anderen Worten, es ist stets möglich, die speziellen Lösungen so zu wählen, dass

$$\xi_{y+\tilde{y}} = \xi_y + \xi_{\tilde{y}} \quad \text{und} \quad \xi_{\lambda y} = \lambda\xi_y$$

gilt, für alle $y, \tilde{y} \in W$ und alle $\lambda \in \mathbb{R}$. Wir werden unten so eine Abbildung ξ für das Gleichungssystem (I.1) angeben, siehe (I.9).

Schließlich wollen wir das Gleichungssystem (I.1) auch explizit lösen und schreiben es dazu nochmals an:

$$\begin{array}{rccccccc} x_1 & -x_2 & -9x_3 & +x_4 & +x_5 & +x_6 & +3x_7 & = & y_1 \\ 2x_1 & -x_2 & -9x_3 & +9x_4 & +3x_5 & +7x_6 & +11x_7 & = & y_2 \\ -x_1 & +3x_2 & +27x_3 & +13x_4 & +2x_5 & +9x_6 & +9x_7 & = & y_3 \\ x_1 & -2x_2 & -18x_3 & -6x_4 & +2x_5 & -4x_6 & +2x_7 & = & y_4 \\ 2x_1 & +x_2 & +9x_3 & +23x_4 & +8x_5 & +17x_6 & +27x_7 & = & y_5 \end{array}$$

¹In der Analysis werden Funktionen der Form $f: \mathbb{R} \rightarrow \mathbb{R}$, $f(x) = ax + b$, als linear bezeichnet. Diese Abbildungen sind nur dann linear im Sinn der linearen Algebra, wenn $b = 0$ gilt, andernfalls ist ja $f(x + \tilde{x}) = a(x + \tilde{x}) + b \neq ax + b + a\tilde{x} + b = f(x) + f(\tilde{x})$. In der linearen Algebra werden solche Funktionen *affin* genannt.

Ist nun y ein Vektor in \mathbb{R}^5 , dessen Komponenten den letzten beiden Gleichungen genügen, dann besitzt das gesamte Gleichungssystem eine Lösung, wir können uns nämlich x_3, x_4, x_6 und x_7 beliebig vorgeben, die anderen Komponenten x_1, x_2 und x_5 sind dann durch die ersten drei Gleichungen eindeutig bestimmt. Wählen wir etwa $x_3 = x_4 = x_6 = x_7 = 0$, erhalten wir die spezielle Lösung

$$\xi_y = \begin{pmatrix} -11y_1 + 5y_2 - 2y_3 \\ -7y_1 + 3y_2 - y_3 \\ 0 \\ 0 \\ 5y_1 - 2y_2 + y_3 \\ 0 \\ 0 \end{pmatrix} = y_1 \begin{pmatrix} -11 \\ -7 \\ 0 \\ 0 \\ 5 \\ 0 \\ 0 \end{pmatrix} + y_2 \begin{pmatrix} 5 \\ 3 \\ 0 \\ 0 \\ -2 \\ 0 \\ 0 \end{pmatrix} + y_3 \begin{pmatrix} -2 \\ -1 \\ 0 \\ 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}. \quad (\text{I.9})$$

Der Teilraum jener $y \in \mathbb{R}^5$, für die das Gleichungssystem eine Lösung besitzt ist daher durch die letzten beiden Gleichungen beschrieben, d.h.

$$W = \left\{ y \in \mathbb{R}^5 \mid \begin{array}{cccc} -13y_1 & +5y_2 & -2y_3 & +y_4 & = 0 \\ -11y_1 & +3y_2 & -3y_3 & & +y_5 = 0 \end{array} \right\}. \quad (\text{I.10})$$

Wollen wir Vektoren y in W finden, so können wir uns die Komponenten y_1, y_2 und y_3 beliebig vorgeben, die beiden Gleichungen in (I.10) bestimmen dann die restlichen Komponenten y_4 und y_5 eindeutig. Etwa sind

$$w_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 13 \\ 11 \end{pmatrix}, \quad w_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ -5 \\ -3 \end{pmatrix}, \quad w_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 2 \\ 3 \end{pmatrix} \quad (\text{I.11})$$

aus W , und jedes Element von $y \in W$ lässt sich in der Form

$$y = t_1 \begin{pmatrix} 1 \\ 0 \\ 0 \\ 13 \\ 11 \end{pmatrix} + t_2 \begin{pmatrix} 0 \\ 1 \\ 0 \\ -5 \\ -3 \end{pmatrix} + t_3 \begin{pmatrix} 0 \\ 0 \\ 1 \\ 2 \\ 3 \end{pmatrix} = \begin{pmatrix} t_1 \\ t_2 \\ t_3 \\ 13t_1 - 5t_2 + 2t_3 \\ 11t_1 - 3t_2 + 3t_3 \end{pmatrix} \quad (\text{I.12})$$

schreiben, wobei die Skalare $t_1, t_2, t_3 \in \mathbb{R}$ eindeutig bestimmt sind. Dies bedeutet gerade, dass die Vektoren w_1, w_2 und w_3 eine Basis von W bilden, es gilt daher

$$\dim(W) = k = 3.$$

Die lineare Parametrisierung $\mathbb{R}^3 \xrightarrow{\cong} W$ ist durch (I.12) gegeben, vgl. (I.6).

Für den Lösungsraum des assoziierten homogenen Gleichungssystems gilt

$$L = \left\{ x \in \mathbb{R}^7 \mid \begin{array}{cccc} x_1 & & +8x_4 & +6x_6 & +4x_7 & = 0 \\ & x_2 & +9x_3 & +7x_4 & +5x_6 & +3x_7 & = 0 \\ & & & & x_5 & +2x_7 & = 0 \end{array} \right\}. \quad (\text{I.13})$$

Wollen wir Vektoren in L finden, können wir uns die Komponenten x_7, x_6, x_4 und x_3 beliebig vorgeben, die restlichen Komponenten sind dann durch die drei Gleichungen bestimmt. So erhalten wir die folgenden Elemente von L

$$b_1 = \begin{pmatrix} -4 \\ -3 \\ 0 \\ 0 \\ -2 \\ 0 \\ 1 \end{pmatrix}, \quad b_2 = \begin{pmatrix} -6 \\ -5 \\ 0 \\ 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \quad b_3 = \begin{pmatrix} -8 \\ -7 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \quad b_4 = \begin{pmatrix} 0 \\ -9 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}. \quad (\text{I.14})$$

Jedes $x \in L$ lässt sich damit in der Form

$$x = s_1 \begin{pmatrix} -4 \\ -3 \\ 0 \\ 0 \\ -2 \\ 0 \\ 1 \end{pmatrix} + s_2 \begin{pmatrix} -6 \\ -5 \\ 0 \\ 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} + s_3 \begin{pmatrix} -8 \\ -7 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} + s_4 \begin{pmatrix} 0 \\ -9 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} -4s_1 - 6s_2 - 8s_3 \\ -3s_1 - 5s_2 - 7s_3 - 9s_4 \\ s_4 \\ s_3 \\ -2s_1 \\ s_2 \\ s_1 \end{pmatrix}$$

schreiben, für eindeutig bestimmte Skalare $s_1, s_2, s_3, s_4 \in \mathbb{R}$. Dies ist also die gesuchte Parametrisierung $\phi: \mathbb{R}^4 \rightarrow L$, siehe (I.5). Die Vektoren b_1, b_2, b_3, b_4 bilden daher eine Basis von L , es ist somit

$$\dim(L) = l = 4.$$

Beachte, dass tatsächlich $k + l = 7$ gilt, vgl. (I.7).

Damit ist das Gleichungssystem (I.1) vollständig gelöst. Die Beschreibung (I.10) von W sagt uns für welche $y \in \mathbb{R}^5$ Lösungen existieren. Ist $y \in W$, d.h. existiert wenigstens eine Lösung, dann lassen sich alle Lösungen x von (I.1) in der Form

$$x = \begin{pmatrix} -11y_1 + 5y_2 - 2y_3 \\ -7y_1 + 3y_2 - y_3 \\ 0 \\ 0 \\ 5y_1 - 2y_2 + y_3 \\ 0 \\ 0 \end{pmatrix} + s_1 \begin{pmatrix} -4 \\ -3 \\ 0 \\ 0 \\ -2 \\ 0 \\ 1 \end{pmatrix} + s_2 \begin{pmatrix} -6 \\ -5 \\ 0 \\ 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} + s_3 \begin{pmatrix} -8 \\ -7 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} + s_4 \begin{pmatrix} 0 \\ -9 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

schreiben, wobei die Skalare $s_1, s_2, s_3, s_4 \in \mathbb{R}$ eindeutig bestimmt sind. Dies ist also die gesuchte Parametrisierung $\phi_y: \mathbb{R}^4 \rightarrow L_y$, vgl. (I.8).

I.2. Geometrische Interpretation. Nach Wahl eines *kartesischen Koordinatensystems* lassen sich Punkte der *Euklidischen Ebene* mit Elementen von \mathbb{R}^2 identifizieren. Der Koordinatenursprung entspricht dabei dem Vektor $0 \in \mathbb{R}^2$. Diese Identifizierung erlaubt es geometrische Konstruktionen und Probleme in

algebraische zu übersetzen. Etwa lässt sich jede Gerade g der Ebene durch eine lineare Gleichung beschreiben, d.h. es existieren $0 \neq (a, b) \in \mathbb{R}^2$ und $c \in \mathbb{R}$, sodass

$$g = \left\{ \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2 : ax + by = c \right\}.$$

Der Schnitt von zwei Geraden entspricht daher der Lösungsmenge eines linearen Gleichungssystems mit zwei Gleichungen in den Unbekannten x und y . Das geometrische Problem den Schnitt zweier Geraden zu bestimmen führt daher auf ein algebraisches Problem, nämlich ein lineares Gleichungssystem, das sich mit den Methoden der linearen Algebra (sehr leicht) lösen lässt.

Darüber hinaus lassen sich einige interessante *geometrische Transformationen* mit linearer Algebra studieren. Ist etwa $\lambda > 0$, dann entspricht die Abbildung

$$\mathbb{R}^2 \rightarrow \mathbb{R}^2, \quad \begin{pmatrix} x \\ y \end{pmatrix} \mapsto \lambda \begin{pmatrix} x \\ y \end{pmatrix},$$

einer beim Koordinatenursprung zentrierten *Streckung* um den Faktor λ . Für fixes $\begin{pmatrix} \tilde{x} \\ \tilde{y} \end{pmatrix} \in \mathbb{R}^2$ entspricht die Abbildung

$$\mathbb{R}^2 \rightarrow \mathbb{R}^2, \quad \begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} \tilde{x} \\ \tilde{y} \end{pmatrix},$$

einer *Translation*. Die Abbildung

$$\rho_\alpha : \mathbb{R}^2 \rightarrow \mathbb{R}^2, \quad \rho_\alpha \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x \cos \alpha - y \sin \alpha \\ x \sin \alpha + y \cos \alpha \end{pmatrix},$$

beschreibt eine beim Koordinatenursprung zentrierte *Drehung* um den Winkel α . Beachte, dass diese Abbildungen ρ linear ist, d.h. es gilt

$$\rho_\alpha \left(\begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} \tilde{x} \\ \tilde{y} \end{pmatrix} \right) = \rho_\alpha \begin{pmatrix} x \\ y \end{pmatrix} + \rho_\alpha \begin{pmatrix} \tilde{x} \\ \tilde{y} \end{pmatrix} \quad \text{und} \quad \rho_\alpha \left(\lambda \begin{pmatrix} x \\ y \end{pmatrix} \right) = \lambda \rho_\alpha \begin{pmatrix} x \\ y \end{pmatrix},$$

für beliebige $\begin{pmatrix} x \\ y \end{pmatrix}, \begin{pmatrix} \tilde{x} \\ \tilde{y} \end{pmatrix} \in \mathbb{R}^2$ und alle $\lambda \in \mathbb{R}$. Die Abbildung

$$\sigma : \mathbb{R}^2 \rightarrow \mathbb{R}^2, \quad \sigma \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x \\ -y \end{pmatrix},$$

entspricht offensichtlich einer *Spiegelung* an der x -Achse. Also beschreibt

$$\sigma_\alpha = \rho_\alpha \circ \sigma \circ \rho_\alpha^{-1} : \mathbb{R}^2 \rightarrow \mathbb{R}^2, \quad \begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} x \cos(2\alpha) + y \sin(2\alpha) \\ x \sin(2\alpha) - y \cos(2\alpha) \end{pmatrix}$$

eine Spiegelung an der Geraden durch den Koordinatenursprung, die mit der x -Achse den Winkel α einschließt. Auch σ_α ist eine lineare Abbildung. Schließlich sei noch erwähnt, dass *Orthogonalprojektionen* auf Geraden durch den Koordinatenursprung ebenfalls durch lineare Abbildungen beschrieben werden können.

Analog lassen sich Punkte des 3-dimensionalen *Euklidischen Raums* nach Wahl eines kartesischen Koordinatensystems mit Elementen von \mathbb{R}^3 identifizieren. Jede Ebene ε im Raum entspricht dabei der Lösungsmenge einer linearen

Gleichung, d.h. es existieren $0 \neq (a, b, c) \in \mathbb{R}^3$ und $d \in \mathbb{R}$, sodass

$$\varepsilon = \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} : ax + by + cz = d \right\}.$$

Der Durchschnitt zweier Ebenen entspricht somit der Lösungsmenge eines linearen Gleichungssystems mit zwei Gleichungen in den Unbekannten x, y, z und lässt sich daher mit den Methoden der linearen Algebra (sehr leicht) bestimmen. Auch können interessante geometrische Transformationen des Euklidischen Raums wie Spiegelungen, Rotationen, Orthogonalprojektionen, u.s.w. durch lineare Abbildungen beschrieben werden.

Wie wir oben gesehen haben, besitzt die lineare Algebra zweifellos Anwendungen in der Euklidischen Geometrie. Wahrscheinlich wichtiger ist jedoch folgender Aspekt. Gleichungssysteme in zwei oder drei Variablen haben geometrische Interpretationen, für die wir eine sehr ausgeprägte Intuition besitzen. Diese Intuition hilft uns, die bei größeren Gleichungssystemen auftretenden Phänomene besser zu begreifen. Wir wollen das noch an einem Beispiel erläutern, und betrachten ein homogenes Gleichungssystem in drei Variablen x, y und z ,

$$a_1x + b_1y + c_1z = 0$$

$$a_2x + b_2y + c_2z = 0$$

mit Koeffizienten $0 \neq (a_1, b_1, c_1) \in \mathbb{R}^3$ und $0 \neq (a_2, b_2, c_2) \in \mathbb{R}^3$. Die geometrische Interpretation der Lösungsmenge als Durchschnitt zweier Ebenen zusammen mit unserer geometrischen Intuition lässt uns vermuten, dass dann wohl eine ganze Gerade in der Lösungsmenge des Systems enthalten sein muss. Dies ist tatsächlich der Fall, formal lässt sich dies aus der weiter oben erwähnten Dimensionsformel herleiten. Allgemeiner folgt aus dieser Dimensionsformel: Jedes homogene Gleichungssystem mit m Gleichungen in n Variablen, wobei $n \geq m$, hat einen mindestens $(n - m)$ -dimensionalen Lösungsraum. Wir können dieses algebraische Resultat zu einem gewissen Grad geometrisch begreifen.

II. Vektorräume und lineare Abbildungen

Gleichungssysteme mit komplexen oder rationalen Koeffizienten lassen sich genau wie reelle Gleichungssysteme lösen. Wichtig dabei ist nur, dass die Koeffizienten in einem Körper liegen. Alle diese Fälle können bequem gleichzeitig behandelt werden. Wir beginnen dieses Kapitel daher mit der Definition von Vektorräumen über beliebigen Körpern \mathbb{K} . Die für uns hier wichtigsten Beispiele werden \mathbb{K}^n und Teilräume davon sein. Anschließend werden wir lineare Abbildungen zwischen Vektorräumen einführen und einige einfache Eigenschaften herleiten. Lineare Abbildungen von \mathbb{K}^n nach \mathbb{K}^m lassen sich effizient durch Matrizen beschreiben, wir werden diesen Zusammenhang in Abschnitt II.4 eingehend diskutieren. In Abschnitt II.5 werden wir besprechen, wie Vektorräume als Summe von Teilräumen zerlegt werden können, und was dies mit (speziellen) Lösungen linearer Gleichungssysteme zu tun hat. Im letzten Abschnitt II.6 werden wir mit den Quotientenräumen die ersten Vektorräume kennen lernen, die nicht in offensichtlicher Weise als Teilräume von Funktionenräumen auftreten.

II.1. Vektorräume. Wir erinnern uns an den Begriff eines Körpers. Eine Menge \mathbb{K} zusammen mit zwei Verknüpfungen $\mathbb{K} \times \mathbb{K} \xrightarrow{+} \mathbb{K}$ und $\mathbb{K} \times \mathbb{K} \xrightarrow{\cdot} \mathbb{K}$ wird Körper genannt, falls die folgenden sogenannten *Körperaxiome* gelten:

- (K1) $\forall \lambda, \mu, \nu \in \mathbb{K} : (\lambda + \mu) + \nu = \lambda + (\mu + \nu)$
- (K2) $\exists 0 \in \mathbb{K} \forall \mu \in \mathbb{K} : \mu + 0 = 0$
- (K3) $\forall \mu \in \mathbb{K} \exists (-\mu) \in \mathbb{K} : \mu + (-\mu) = 0$
- (K4) $\forall \lambda, \mu \in \mathbb{K} : \lambda + \mu = \mu + \lambda$
- (K5) $\forall \lambda, \mu, \nu \in \mathbb{K} : (\lambda\mu)\nu = \lambda(\mu\nu)$
- (K6) $\exists 1 \in \mathbb{K} : 1 \neq 0 \wedge \forall \lambda \in \mathbb{K} : \lambda 1 = \lambda$
- (K7) $\forall \lambda \in \mathbb{K} \setminus \{0\} \exists \lambda^{-1} \in \mathbb{K} : \lambda \lambda^{-1} = 1$
- (K8) $\forall \lambda, \mu \in \mathbb{K} : \lambda\mu = \mu\lambda$
- (K9) $\forall \lambda, \mu, \nu \in \mathbb{K} : \lambda(\mu + \nu) = \lambda\mu + \lambda\nu$

Ist \mathbb{K} ein Körper, dann bildet \mathbb{K} bezüglich der *Addition* eine abelsche Gruppe mit neutralem Element 0. Darüber hinaus ist auch $\mathbb{K} \setminus \{0\}$ bezüglich der *Multiplikation* eine abelsche Gruppe mit neutralem Element 1. Das Distributivgesetz (K9) garantiert eine gewisse Verträglichkeit zwischen Addition und Multiplikation. Einfache Konsequenzen aus den Körperaxiomen, wie z.B. $\forall \lambda \in \mathbb{K} : 0\lambda = 0$, die Eindeutigkeit der beiden neutralen Elemente 0 und 1, oder die Nullteilerfreiheit, werden wir im Folgenden als bekannt voraussetzen, siehe etwa [7, Kapitel 5]. Auch werden wir üblichen Konventionen folgen und schreiben etwa $\lambda\mu\nu$ statt $(\lambda\mu)\nu = \lambda(\mu\nu)$ und manchmal $\overset{\mu}{\lambda}$ für $\lambda\mu^{-1}$.

Als wichtige Beispiele von Körpern sind jedenfalls \mathbb{R} , \mathbb{C} , \mathbb{Q} und \mathbb{Z}_p zu nennen, wobei p eine Primzahl bezeichnet. Beachte, dass \mathbb{Z}_p nur aus endlich vielen Elementen besteht. Der kleinste Körper, \mathbb{Z}_2 , besteht überhaupt nur aus 0 und 1.

Ist \mathbb{K} ein Körper, dann existiert ein eindeutiger Ringhomomorphismus $\mathbb{Z} \rightarrow \mathbb{K}$, der $1 \in \mathbb{Z}$ auf $1 \in \mathbb{K}$ abbildet. Jede ganze Zahl $n \in \mathbb{Z}$ können wir daher auch

als Element von \mathbb{K} auffassen, etwa gilt $2 = 1 + 1$, $3 = 1 + 1 + 1$, u.s.w. Beachte jedoch, dass in manchen Körpern und für manche $n \in \mathbb{N}$ sehr wohl $n = 0 \in \mathbb{K}$ gelten kann, etwa haben wir im Körper \mathbb{Z}_2 die Relation $2 = 1 + 1 = 0$.

II.1.1. DEFINITION (Vektorraum). Unter einem *Vektorraum* über einem Körper \mathbb{K} verstehen wir eine Menge V zusammen mit zwei Verknüpfungen,

$$V \times V \xrightarrow{+} V \quad \text{und} \quad \mathbb{K} \times V \xrightarrow{\cdot} V,$$

die folgenden acht Axiomen, den sogenannten *Vektorraumaxiomen*, genügen:

$$(V1) \quad \forall u, v, w \in V : (u + v) + w = u + (v + w)$$

$$(V2) \quad \exists 0 \in V \forall v \in V : v + 0 = v$$

$$(V3) \quad \forall v \in V \exists (-v) \in V : v + (-v) = 0$$

$$(V4) \quad \forall v, w \in V : v + w = w + v$$

$$(V5) \quad \forall \lambda, \mu \in \mathbb{K} \forall v \in V : \lambda(\mu v) = (\lambda\mu)v$$

$$(V6) \quad \forall \lambda \in \mathbb{K} \forall v, w \in V : \lambda(v + w) = \lambda v + \lambda w$$

$$(V7) \quad \forall \lambda, \mu \in \mathbb{K} \forall v \in V : (\lambda + \mu)v = \lambda v + \mu v$$

$$(V8) \quad \forall v \in V : 1v = v$$

Die beiden Verknüpfungen werden als (*Vektor-*)*Addition* bzw. *Skalarmultiplikation* bezeichnet. Elemente eines Vektorraums werden *Vektoren* genannt. Ein Vektorraum über \mathbb{K} wird oft auch als \mathbb{K} -Vektorraum bezeichnet. Ist der Grundkörper $\mathbb{K} = \mathbb{R}$, so sprechen wir von einem *reellen Vektorraum*, im Fall $\mathbb{K} = \mathbb{C}$ von einem *komplexen Vektorraum*.

Es sei V ein Vektorraum über \mathbb{K} . Aufgrund von Axiom (V1) können wir bei mehrfachen Summen auf die Klammersetzung verzichten und schreiben meist $u + v + w$, $v_1 + \dots + v_n$ oder $\sum_{i=1}^n v_i$.

Nach Axiom (V2) existiert $0 \in V$, sodass $v + 0 = v$ für jedes $v \in V$. Der Vektor 0 ist durch diese Eigenschaft eindeutig bestimmt, ist nämlich $o \in V$ ein weiteres Element mit $v + o = v$ für alle $v \in V$, so erhalten wir mit Hilfe von Axiom (V4) sofort $0 = 0 + o = o + 0 = o$. Dieses additiv neutrale Element 0 wird *Nullvektor* genannt. Nach Axiom (V4) gilt jedes $v \in V$ auch $v + 0 = v = 0 + v$.

Nach Axiom (V3) existiert zu jedem $v \in V$ ein Element $-v \in V$, sodass $v + (-v) = 0$. Dieser Vektor $-v$ ist eindeutig bestimmt, ist nämlich $v' \in V$ ein weiteres Element mit $v + v' = 0$, dann folgt mit Hilfe von Axiom (V4) sofort $v' = v' + 0 = v' + (v + (-v)) = (v' + v) + (-v) = (v + v') + (-v) = 0 + (-v) = -v$. Nach Axiom (V4) gilt für beliebiges $v \in V$ auch $v + (-v) = 0 = (-v) + v$. Weiters ist $-0 = 0$, denn $0 + 0 = 0$. Für $v, w \in V$ haben wir $-(v + w) = (-v) + (-w)$, denn $(v + w) + ((-v) + (-w)) = (v + (-v)) + (w + (-w)) = 0 + 0 = 0$. Beachte auch $-(-v) = v$.

In jedem Vektorraum gilt die Kürzungsregel

$$\forall u, v, w \in V : u + v = w + v \Rightarrow u = w,$$

denn aus $u + v = w + v$ folgt $u = u + 0 = u + (v + (-v)) = (u + v) + (-v) = (w + v) + (-v) = w + (v + (-v)) = w + 0 = w$.

Bis jetzt haben wir nur die Axiome (V1) bis (V4) verwendet, die in den vorangehenden Absätzen besprochenen Eigenschaften gelten daher in allgemeinen abelschen Gruppen, siehe [7, Abschnitt 5.2]. Wir widmen uns nun dem Zusammenspiel von Addition und Skalarmultiplikation.

Für jedes $\lambda \in \mathbb{K}$ gilt

$$\lambda 0 = 0,$$

denn aus Axiom (V6) folgt $0 + \lambda 0 = \lambda 0 = \lambda(0 + 0) = \lambda 0 + \lambda 0$ und wegen oben erwähnter Kürzungsregel daher $0 = \lambda 0$.

Für jedes $v \in V$ gilt

$$0v = 0,$$

denn aus Axiom (V7) erhalten wir $0 + 0v = 0v = (0 + 0)v = 0v + 0v$ und wegen der Kürzungsregel oben daher $0 = 0v$.

Es gilt auch folgende Umkehrung der vorangehenden beiden Aussagen:

$$\forall \lambda \in \mathbb{K} \forall v \in V : \lambda v = 0 \Rightarrow \lambda = 0 \text{ oder } v = 0.$$

Ist nämlich $\lambda v = 0$ und $\lambda \neq 0$, dann erhalten wir mit Hilfe der Axiome (V5) und (V8) sofort $v = 1v = (\lambda^{-1}\lambda)v = \lambda^{-1}(\lambda v) = \lambda^{-1}0 = 0$.

Für beliebiges $v \in V$ gilt

$$-v = (-1)v,$$

denn nach (V7) und (V8) ist $v + (-1)v = 1v + (-1)v = (1 + (-1))v = 0v = 0$.

Schließlich haben wir auch

$$\lambda(-v) = -(\lambda v) = (-\lambda)v,$$

für alle $\lambda \in \mathbb{K}$ und $v \in V$, denn mit Hilfe von Axiom (V5) folgt $\lambda(-v) = \lambda((-1)v) = (\lambda(-1))v = (-\lambda)v = ((-1)\lambda)v = (-1)(\lambda v) = -\lambda v$.

Um die Notation zu vereinfachen, schreiben wir von nun an $v - w$ für $v + (-w)$. Auch verzichten wir bei Skalarmultiplikationen oft auf Klammersetzung und schreiben meist $\lambda\mu v$ statt $\lambda(\mu v) = (\lambda\mu)v$, falls $\lambda, \mu \in \mathbb{K}$ und $v \in V$, vgl. Axiom (V5). Nach obigen Bemerkungen ist dies mit den vertrauten Rechenregeln verträglich.

II.1.2. BEISPIEL. Wir können jeden Körper \mathbb{K} als \mathbb{K} -Vektorraum auffassen.

II.1.3. BEISPIEL. Es sei \mathbb{K} ein Körper und $n \in \mathbb{N}$. Auf der Menge aller n -Tupel von Elementen aus \mathbb{K} ,

$$\mathbb{K}^n := \left\{ \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} : x_1, \dots, x_n \in \mathbb{K} \right\},$$

definieren wir Addition, $\mathbb{K}^n \times \mathbb{K}^n \xrightarrow{+} \mathbb{K}^n$, und Skalarmultiplikation, $\mathbb{K} \times \mathbb{K}^n \xrightarrow{\cdot} \mathbb{K}^n$, komponentenweise, d.h. durch

$$\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} + \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} := \begin{pmatrix} x_1 + y_1 \\ \vdots \\ x_n + y_n \end{pmatrix} \quad \text{und} \quad \lambda \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} := \begin{pmatrix} \lambda x_1 \\ \vdots \\ \lambda x_n \end{pmatrix}$$

wobei $\lambda \in \mathbb{K}$. Mit diesen Operationen wird \mathbb{K}^n zu einem Vektorraum über \mathbb{K} , die Vektorraumaxiome folgen sofort aus den entsprechenden Körperaxiomen, siehe Übungsaufgabe 11 oder [7, Kapitel 7]. Insbesondere ist \mathbb{R}^n ein reeller Vektorraum und \mathbb{C}^n ein komplexer Vektorraum. Etwa gilt in \mathbb{C}^3 ,

$$\begin{pmatrix} 1 + 2\mathbf{i} \\ 3 - 4\mathbf{i} \\ \mathbf{i} \end{pmatrix} + \begin{pmatrix} 5 \\ 6\mathbf{i} \\ 2 - \mathbf{i} \end{pmatrix} = \begin{pmatrix} 6 + 2\mathbf{i} \\ 3 + 2\mathbf{i} \\ 2 \end{pmatrix} \quad \text{und} \quad (2 + \mathbf{i}) \begin{pmatrix} 4 + 3\mathbf{i} \\ -2\mathbf{i} \\ 7 \end{pmatrix} = \begin{pmatrix} 5 + 10\mathbf{i} \\ 2 - 4\mathbf{i} \\ 14 + 7\mathbf{i} \end{pmatrix}.$$

II.1.4. BEMERKUNG. Die Wahl eines *kartesischen Koordinatensystems* ermöglicht es Elemente von \mathbb{R} , \mathbb{R}^2 bzw. \mathbb{R}^3 mit Punkten der *Gerade*, der *Ebene* bzw. des *Raums* zu identifizieren, siehe [7, Kapitel 7]. In diesem Bild besitzen die Vektorraumoperationen geometrische Interpretationen. Skalarmultiplikation mit $0 \neq \lambda \in \mathbb{R}$, d.h. die Abbildung $x \mapsto \lambda x$, entspricht einer *Streckung* um den Faktor λ , zentriert beim Ursprung des Koordinatensystems. Für fixes y entspricht die Abbildung $x \mapsto x + y$ einer *Translation* um y .

II.1.5. BEISPIEL. Über jedem Körper \mathbb{K} gibt es einen Vektorraum der nur aus einem Element, dem Nullvektor, besteht. Wir bezeichnen diesen *triviale Vektorraum* mit $\{0\}$, \mathbb{K}^0 oder 0 .

II.1.6. BEISPIEL (Funktionenräume). Es sei X eine Menge und V ein Vektorraum über einem Körper \mathbb{K} . Es bezeichne $F(X, V) = V^X$ die Menge aller Abbildungen $X \rightarrow V$. Die Summe zweier Abbildungen $f, g \in F(X, V)$ wird punktweise definiert, d.h.

$$f + g: X \rightarrow V, \quad (f + g)(x) := f(x) + g(x).$$

Für $\lambda \in \mathbb{K}$ und $f \in F(X, V)$ definieren wir analog

$$\lambda f: X \rightarrow V, \quad (\lambda f)(x) := \lambda f(x).$$

Mit diesen Operationen wird $F(X, V)$ zu einem \mathbb{K} -Vektorraum. Etwa haben wir für $f, g, h \in F(X, V)$ und $x \in X$

$$\begin{aligned} ((f + g) + h)(x) &= (f + g)(x) + h(x) = (f(x) + g(x)) + h(x) \\ &= f(x) + (g(x) + h(x)) = f(x) + (g + h)(x) = (f + (g + h))(x). \end{aligned}$$

Da dies für alle $x \in X$ gilt, erhalten wir $(f + g) + h = f + (g + h)$, also genügt $F(X, V)$ dem Vektorraumaxiom (V1). Auch Axiom (V2) ist erfüllt, die konstante Nullfunktion, $0: X \rightarrow V$, $0(x) := 0$, ist neutrales Element der Addition, denn für jedes $f \in F(X, V)$ und $x \in X$ gilt $(f + 0)(x) = f(x) + 0(x) = f(x) + 0 = f(x)$,

also $f + 0 = f$. Das additive Inverse von $f \in F(X, V)$ ist durch $-f: X \rightarrow V$, $(-f)(x) := -f(x)$, gegeben, denn für jedes $x \in X$ gilt $(f + (-f))(x) = f(x) + (-f)(x) = f(x) - f(x) = 0$, also $f + (-f) = 0$. Damit ist auch Axiom (V3) für $F(X, V)$ verifiziert. Für $\lambda \in \mathbb{K}$, $f, g \in F(X, V)$ und $x \in X$ haben wir

$$\begin{aligned} (\lambda(f + g))(x) &= \lambda(f + g)(x) = \lambda(f(x) + g(x)) \\ &= \lambda f(x) + \lambda g(x) = (\lambda f)(x) + (\lambda g)(x) = (\lambda f + \lambda g)(x). \end{aligned}$$

Da dies für alle $x \in X$ gilt, erhalten wir $\lambda(f + g) = \lambda f + \lambda g$, d.h. $F(X, V)$ genügt Axiom (V6). Völlig analog lassen sich die verbleibenden Vektorraumaxiome überprüfen, siehe Übungsaufgabe 13.

Wählen wir speziell $V = \mathbb{K}$, so sehen wir, dass $F(X; \mathbb{K})$, d.h. die Menge aller \mathbb{K} -wertigen Funktionen auf X , bezüglich punktweiser Addition und Skalarmultiplikation einen \mathbb{K} -Vektorraum bilden. Etwa bilden die reellwertigen Funktionen auf einem Intervall, $F([a, b], \mathbb{R})$, einen reellen Vektorraum. Für $X = \mathbb{N}$ ist $F(\mathbb{N}, \mathbb{R})$ die Menge aller Folgen in \mathbb{R} , diese bilden daher bezüglich gliedweiser Addition und Skalarmultiplikation ebenfalls einen Vektorraum.

II.1.7. BEMERKUNG (Funktionen als Algebra). \mathbb{K} -wertige Funktionen können auch in naheliegender Weise multipliziert werden, für $f, g \in F(X, \mathbb{K})$ wird ihr Produkt $fg \in F(X, \mathbb{K})$ punktweise, d.h. durch

$$fg: X \rightarrow \mathbb{K}, \quad (fg)(x) := f(x)g(x),$$

definiert, $x \in X$. Diese Verknüpfung ist assoziativ und kommutativ, d.h. es gilt $f(gh) = (fg)h$ sowie $fg = gf$ für beliebige $f, g, h \in F(X, \mathbb{K})$. Die konstante Einsfunktion, $1 \in F(X, \mathbb{K})$, ist neutrales Element bezüglich der Multiplikation, d.h. für alle $f \in F(X, \mathbb{K})$ gilt $1f = f = f1$. Die Multiplikation ist mit der Vektorraumstruktur verträglich, es gilt $f(g_1 + g_2) = fg_1 + fg_2$, $(f_1 + f_2)g = f_1g + f_2g$ und $f(\lambda g) = \lambda(fg) = (\lambda f)g$, für beliebige $f, f_1, f_2, g, g_1, g_2 \in F(X; \mathbb{K})$ und $\lambda \in \mathbb{K}$. Dies bedeutet gerade, dass $F(X, \mathbb{K})$ eine kommutative \mathbb{K} -Algebra mit Eins² bildet, vgl. Übungsaufgabe 14.

II.1.8. BEISPIEL (Polynome). Sei \mathbb{K} ein Körper. Unter einem *Polynom mit Koeffizienten in \mathbb{K}* verstehen wir einen formalen Ausdruck der Form

$$p_0 + p_1z + p_2z^2 + p_3z^3 + \dots$$

wobei die Koeffizienten $p_i \in \mathbb{K}$ fast alle verschwinden, d.h. alle bis auf endlich viele p_i gleich 0 sind. Zwei Polynome werden als gleich betrachtet, wenn alle ihre

²Unter einer \mathbb{K} -Algebra verstehen wir einen \mathbb{K} -Vektorraum A zusammen mit einer Verknüpfung $A \times A \rightarrow A$, der sogenannten Multiplikation, sodass $a(b_1 + b_2) = ab_1 + ab_2$, $a(\lambda b) = \lambda(ab)$, $(a_1 + a_2)b = a_1b + a_2b$ und $(\lambda a)b = \lambda(ab)$, für alle $a, a_1, a_2, b, b_1, b_2 \in A$ und $\lambda \in \mathbb{K}$. Die Algebra wird assoziativ genannt, wenn $a(bc) = (ab)c$ für alle $a, b, c \in A$ gilt. Ein Element $e \in A$ mit $ae = a = ea$ für alle $a \in A$, wird Einselement von A genannt. Dieses neutrale Element der Multiplikation ist dann eindeutig bestimmt und wir sprechen von einer assoziativen Algebra mit Eins. Gilt darüber hinaus $ab = ba$ für alle $a, b \in A$, dann wird A kommutativ genannt.

Koeffizienten überein stimmen. Die Menge aller Polynome mit Koeffizienten in \mathbb{K} bezeichnen wir mit $\mathbb{K}[z]$. Sind $p, q \in \mathbb{K}[z]$ zwei Polynome, $p = p_0 + p_1z + p_2z^2 + \dots$ und $q = q_0 + q_1z + q_2z^2 + \dots$, so wird ihre Summe koeffizientenweise definiert,

$$p + q := (p_0 + q_0) + (p_1 + q_1)z + (p_2 + q_2)z^2 + (p_3 + q_3)z^3 + \dots$$

Offensichtlich ist dann $p + q$ wieder ein Polynom, d.h. $p + q \in \mathbb{K}[z]$. Analog definieren wir Skalarmultiplikation mit $\lambda \in \mathbb{K}$ durch

$$\lambda p := (\lambda p_0) + (\lambda p_1)z + (\lambda p_2)z^2 + (\lambda p_3)z^3 + \dots$$

und erhalten ein Polynom $\lambda p \in \mathbb{K}[z]$. Mit diesen Verknüpfungen wird $\mathbb{K}[z]$ zu einem Vektorraum über \mathbb{K} . Dabei ist das Nullpolynom, $0 := 0 + 0z + 0z^2 + \dots$, das additiv neutrale Element, d.h. $0 + p = p = p + 0$ für alle $p \in \mathbb{K}[z]$. Meist werden bei Polynomen nur jene Potenzen von z angeschrieben, deren Koeffizienten verschieden von 0 sind. Etwa sind $p = 2 - z + 4z^2 - 7z^5$ und $q = z - z^3$ zwei reelle Polynome für die $p + q = 2 + 4z^2 - z^3 - 7z^5$ und $7q = 7z - 7z^3$ gilt.

II.1.9. BEMERKUNG (Polynome als Algebra). Auch Polynome können multipliziert werden. Sind $p = \sum_i p_i z^i = p_0 + p_1z + p_2z^2 + \dots$ und $q = \sum_j q_j z^j = q_0 + q_1z + q_2z^2 + \dots$ zwei Polynome in $\mathbb{K}[z]$, so wird ihr Produkt $pq \in \mathbb{K}[z]$ durch

$$pq = p_0p_0 + (p_0q_1 + p_1q_0)z + (p_0q_2 + p_1q_1 + p_2q_0)z^2 + \dots + \left(\sum_{i+j=k} p_i q_j \right) z^k$$

d.h. durch *formales Ausmultiplizieren* definiert. Etwa ist das Produkt der Polynome $p = 2 + 3z + 4z^2$ und $q = 1 + z - z^2$ gleich $pq = 2 + 5z + 5z^2 + z^3 - 4z^4$. Die Multiplikation von Polynomen ist assoziativ, d.h. für $p, q, r \in \mathbb{K}[z]$ gilt $p(qr) = (pq)r$, denn

$$\begin{aligned} p(qr) &= \left(\sum_i p_i z^i \right) \left(\left(\sum_j q_j z^j \right) \left(\sum_k r_k z^k \right) \right) \\ &= \left(\sum_i p_i z^i \right) \left(\sum_l \left(\sum_{j+k=l} q_j r_k \right) z^l \right) \\ &= \sum_m \left(\sum_{i+l=m} p_i \sum_{j+k=l} q_j r_k \right) z^m = \sum_m \left(\sum_{i+j+k=m} p_i q_j r_k \right) z^m \end{aligned}$$

und dies stimmt mit dem Polynom

$$\begin{aligned} (pq)r &= \left(\left(\sum_i p_i z^i \right) \left(\sum_j q_j z^j \right) \right) \left(\sum_k r_k z^k \right) \\ &= \left(\left(\sum_{i+j=l} p_i q_j \right) z^l \right) \left(\sum_k r_k z^k \right) \\ &= \sum_m \left(\sum_{l+k=m} \left(\sum_{i+j=l} p_i q_j \right) r_k \right) z^m = \sum_m \left(\sum_{i+j+k=m} p_i q_j r_k \right) z^m \end{aligned}$$

überein. Für $p, q \in \mathbb{K}[z]$ gilt auch $pq = qp$, d.h. die Multiplikation von Polynomen ist kommutativ. Das Einspolynom, $1 = 1 + 0z + 0z^2 + \dots$ ist neutrales Element der Multiplikation, d.h. $1p = p = p1$ für alle $p \in \mathbb{K}[z]$. Schließlich ist die Multiplikation von Polynomen auch mit der Vektorraumstruktur auf $\mathbb{K}[z]$ verträglich, es gilt $r(p + q) = rp + rq$, $(p + q)r = pr + qr$ sowie $p(\lambda q) = \lambda(pq) = (\lambda p)q$, für beliebige Polynome $p, q, r \in \mathbb{K}[z]$ und $\lambda \in \mathbb{K}$. Dies bedeutet gerade, dass $\mathbb{K}[z]$ eine kommutative \mathbb{K} -Algebra mit Eins bildet, vgl. Übungsaufgabe 15.

II.2. Teilräume. Suchen wir nach Teilmengen eines Vektorraums, die selbst einen Vektorraum bilden, werden auf den Begriff des Teilraums geführt.

II.2.1. DEFINITION (Teilraum). Es sei V ein Vektorraum über \mathbb{K} . Eine nicht leere Teilmenge $W \subseteq V$ wird *Teilraum* von V genannt, wenn sie abgeschlossen unter Addition und Skalarmultiplikation ist. In anderen Worten, eine nicht leere Teilmenge $W \subseteq V$ ist genau dann Teilraum von V , wenn sie folgenden beiden Bedingungen genügt:

- (a) $\forall w_1, w_2 \in W : w_1 + w_2 \in W$
- (b) $\forall \lambda \in \mathbb{K} \forall w \in W : \lambda w \in W$

II.2.2. BEMERKUNG. Ist W ein Teilraum von V , dann gilt $0 \in W$. Nach Definition ist W nämlich nicht leer, also existiert $w \in W$. Aus (b) folgt daher $0w \in W$. Da $0w = 0$, erhalten wir somit auch $0 \in W$.

II.2.3. BEMERKUNG. Ist W ein Teilraum von V und $w \in W$, dann gilt auch $-w \in W$. Nach (b) gilt nämlich $(-1)w \in W$, und da $(-1)w = -w$ folgt $-w \in W$.

Ist W ein Teilraum von V , dann schränken sich Addition und Skalarmultiplikation in V zu Abbildungen $W \times W \xrightarrow{+} W$ und $\mathbb{K} \times W \xrightarrow{\cdot} W$ ein. Dadurch wird W selbst zu einem Vektorraum. Die Gültigkeit der Axiome (V1) und (V4) – (V8) für V impliziert sofort, dass diese Axiome auch für W gelten. Nach Bemerkung II.2.2 ist auch (V2) für W erfüllt. Nach Bemerkung II.2.3 genügt W schließlich auch Axiome (V3). Wir halten dies in folgender Proposition fest.

II.2.4. PROPOSITION (Teilräume als Vektorräume). *Sei V ein \mathbb{K} -Vektorraum und W ein Teilraum von V . Dann bildet W bezüglich der eingeschränkten Verknüpfungen selbst einen \mathbb{K} -Vektorraum.*

II.2.5. BEMERKUNG. $\{0\}$ und V sind stets Teilräume von V .

II.2.6. PROPOSITION (Durchschnitt von Teilräumen). *Ist W_i , $i \in I$, eine Familie von Teilräumen eines Vektorraums V , so bildet auch deren Durchschnitt, $\bigcap_{i \in I} W_i$, einen Teilraum von V .*

BEWEIS. Zunächst ist der Durchschnitt jedenfalls nicht leer, denn nach Bemerkung II.2.2 gilt $0 \in \bigcap_{i \in I} W_i$. Um die Abgeschlossenheit unter der Addition zu überprüfen seien nun $w, w' \in \bigcap_{i \in I} W_i$. Für jedes $i \in I$ gilt daher $w, w' \in W_i$. Da W_i einen Teilraum bildet folgt $w + w' \in W_i$. Somit ist $w + w' \in \bigcap_{i \in I} W_i$,

und der Durchschnitt daher abgeschlossen unter Addition. Um die Abgeschlossenheit unter Skalarmultiplikation zu zeigen seien nun $\lambda \in \mathbb{K}$ und $w \in \bigcap_{i \in I} W_i$. Für jedes $i \in I$ gilt daher $w \in W_i$. Da W_i einen Teilraum bildet folgt $\lambda w \in W_i$. Somit ist $\lambda w \in \bigcap_{i \in I} W_i$ und der Durchschnitt daher auch abgeschlossen unter Skalarmultiplikation. \square

II.2.7. BEISPIEL. Sind $n, m \in \mathbb{N}$ und $a_{ij} \in \mathbb{K}$, dann bildet die Menge

$$\left\{ \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{K}^n \mid \begin{array}{l} a_{11}x_1 + \cdots + a_{1n}x_n = 0 \\ \vdots \\ a_{m1}x_1 + \cdots + a_{mn}x_n = 0 \end{array} \right\}$$

einen Teilraum von \mathbb{K}^n . Die Lösungsmenge eines homogenen Gleichungssystems ist daher stets ein Teilraum. Wir werden später sehen, dass sich jeder Teilraum von \mathbb{K}^n in dieser Form darstellen lässt, für geeignete $m \in \mathbb{N}_0$ und $a_{ij} \in \mathbb{K}$. Auch

$$\left\{ x_1 \begin{pmatrix} a_{11} \\ \vdots \\ a_{m1} \end{pmatrix} + \cdots + x_n \begin{pmatrix} a_{1n} \\ \vdots \\ a_{mn} \end{pmatrix} \mid x_1, \dots, x_n \in \mathbb{K} \right\},$$

d.h. die Teilmenge aller Vektoren $y \in \mathbb{K}^m$, für die das Gleichungssystem

$$\begin{array}{rcl} a_{11}x_1 + \cdots + a_{1n}x_n & = & y_1 \\ & \vdots & \\ a_{m1}x_1 + \cdots + a_{mn}x_n & = & y_m \end{array}$$

wenigstens eine Lösung besitzt, bildet einen Teilraum von \mathbb{K}^m . Wir werden später sehen, dass sich jeder Teilraum von \mathbb{K}^m in dieser Form beschreiben lässt, für geeignete $n \in \mathbb{N}_0$ und $a_{ij} \in \mathbb{K}$.

II.2.8. BEISPIEL (Teilräume von \mathbb{K}). Der Vektorraum $\mathbb{K} = \mathbb{K}^1$ besitzt außer den beiden trivialen Teilräumen $\{0\}$ und \mathbb{K} keine weiteren Teilräume. Ist nämlich $W \subseteq \mathbb{K}$ ein Teilraum und $W \neq \{0\}$, dann existiert $w \in W$ mit $0 \neq w$, also $\lambda = (\lambda w^{-1})w \in W$ für jedes $\lambda \in \mathbb{K}$, und daher $W = \mathbb{K}$.

II.2.9. BEISPIEL (Teilräume von \mathbb{K}^2). Neben den trivialen Teilräumen $\{0\}$ und \mathbb{K}^2 bildet auch jede Teilmenge der Form

$$\langle a \rangle = \left\{ \lambda \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} \mid \lambda \in \mathbb{K} \right\} = \left\{ \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \mid a_2x_1 - a_1x_2 = 0 \right\},$$

wobei $0 \neq a = \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} \in \mathbb{K}^2$, einen Teilraum von \mathbb{K}^2 . Wir wollen uns nun überlegen, dass dies schon alle Teilräume von \mathbb{K}^2 sind. Sei dazu $W \subseteq \mathbb{K}^2$ ein beliebiger Teilraum und $\{0\} \neq W$. Dann existiert $0 \neq a \in W$ und wegen der Abgeschlossenheit von W unter Skalarmultiplikation erhalten wir $\langle a \rangle \subseteq W$. Ist $\langle a \rangle \neq W$, dann existiert also $b \in W \setminus \langle a \rangle$. Wir bezeichnen die Koordinaten von a und b mit $a = \begin{pmatrix} a_1 \\ a_2 \end{pmatrix}$ und $b = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}$, $a_1, a_2, b_1, b_2 \in \mathbb{K}$. Aus $a \neq 0$ und $b \notin \langle a \rangle$ folgt $d := a_1b_2 - a_2b_1 \neq 0$.

Da W abgeschlossen unter Addition und Skalarmultiplikation ist, erhalten wir für jeden Vektor $\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \in \mathbb{K}^2$,

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \frac{x_1 b_2 - x_2 b_1}{d} \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} + \frac{x_2 a_1 - x_1 a_2}{d} \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} \in W,$$

und somit $W = \mathbb{K}^2$. Insbesondere hat \mathbb{R}^2 neben den beiden trivialen Teilräumen $\{0\}$ und \mathbb{R}^2 nur Teilräume der Form $\{\lambda a \mid \lambda \in \mathbb{R}\}$ wobei $0 \neq a \in \mathbb{R}^2$. Nach Wahl eines kartesischen Koordinatensystems entsprechen letztere genau den Geraden durch den Ursprung des Koordinatensystems.

II.2.10. BEISPIEL (Teilräume von \mathbb{K}^3). Neben den trivialen Teilräumen $\{0\}$ und \mathbb{K}^3 besitzt der Vektorraum \mathbb{K}^3 auch Teilräume der Form $\langle v \rangle = \{\lambda v \mid \lambda \in \mathbb{K}\}$, $v \in \mathbb{K}^3$, sowie

$$\left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{K}^3 \mid ax + by + cz = 0 \right\},$$

wobei $a, b, c \in \mathbb{K}$. Wir werden später sehen, dass sich jeder Teilraum von \mathbb{K}^3 so beschreiben lässt. Insbesondere hat \mathbb{R}^3 nur Teilräume dieser Gestalt. Nach Wahl eines kartesischen Koordinatensystems entsprechen die nicht-trivialen Fälle genau den Geraden bzw. Ebenen durch den Ursprung des Koordinatensystems.

II.2.11. BEISPIEL. Die Teilmenge $\{ \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \in \mathbb{K}^2 : x_1 x_2 = 0 \}$ von \mathbb{K}^2 ist zwar abgeschlossen unter Skalarmultiplikation bildet jedoch keinen Teilraum. Auch die Teilmenge $\{ \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \in \mathbb{R}^2 : x_1 \geq 0 \}$ bildet keinen Teilraum von \mathbb{R}^2 , obwohl sie abgeschlossen unter Addition ist. Dasselbe gilt für $\mathbb{Z}^2 \subseteq \mathbb{R}^2$. Fassen wir \mathbb{R} als Teilmenge von \mathbb{C} auf, dann ist auch $\mathbb{R}^n \subseteq \mathbb{C}^n$. Obwohl diese Teilmenge abgeschlossen unter Addition ist, bildet sie keinen Teilraum des komplexen Vektorraums \mathbb{C}^n , für $0 \neq x \in \mathbb{R}^n$ gilt nämlich $ix \notin \mathbb{R}^n$.

II.2.12. BEISPIEL (Vektorräume von Funktionen aus der Analysis). Die Menge der reellwertigen stetigen Funktionen auf einem Intervall, $C([a, b], \mathbb{R})$, bildet einen Teilraum von $F([a, b], \mathbb{R})$, denn Summen und skalare Vielfache stetiger Funktionen sind wieder stetig. Ebenso ist die Menge der beschränkten Funktionen $[a, b] \rightarrow \mathbb{R}$ ein Teilraum von $F([a, b], \mathbb{R})$. Auch die (Riemann-)integrierbaren Funktionen $[a, b] \rightarrow \mathbb{R}$ bilden einen Teilraum von $F([a, b], \mathbb{R})$. Genauso bilden die differenzierbaren Abbildungen $(a, b) \rightarrow \mathbb{R}$ einen Teilraum von $F((a, b), \mathbb{R})$. Analog ist die Menge aller k mal stetig differenzierbaren Funktionen, $C^k((a, b), \mathbb{R})$, ein Teilraum von $F((a, b), \mathbb{R})$. Auch die Menge aller glatten Funktionen $C^\infty((a, b), \mathbb{R}) = \bigcap_{k \in \mathbb{N}} C^k((a, b), \mathbb{R})$ ist ein Teilraum von $F((a, b), \mathbb{R})$, vgl. Proposition II.2.6. Nach Proposition II.2.4 sind dies daher alles Vektorräume bezüglich punktweiser Addition und Skalarmultiplikation von Funktionen.

II.2.13. BEISPIEL (Vektorräume von Folgen aus der Analysis). Die Menge der konvergenten Folgen bildet einen Teilraum von $F(\mathbb{N}, \mathbb{R})$, denn Summen und skalare Vielfache konvergenter Folgen sind wieder konvergent. Auch die Menge der

beschränkten Folgen ist ein Teilraum von $F(\mathbb{N}, \mathbb{R})$. Ebenso bilden die summierbaren Folgen (Reihen) einen Teilraum von $F(\mathbb{N}, \mathbb{R})$. Gleiches gilt für die Menge der absolut summierbaren Folgen oder die Menge der quadratsummierbaren Folgen.

II.2.14. BEISPIEL. Es bezeichne $X \subseteq \mathbb{R}$ ein um Null symmetrisches Intervall, etwa $X = (-a, a)$ mit $0 < a \leq \infty$. Dann bildet die Menge der *geraden Funktionen*,

$$G := \{f: X \rightarrow \mathbb{R} \mid \forall x \in \mathbb{R} : f(-x) = f(x)\},$$

einen Teilraum von $F(X, \mathbb{R})$. Für $f, g \in G$ und $x \in X$ gilt nämlich $(f+g)(-x) = f(-x) + g(-x) = f(x) + g(x) = (f+g)(x)$, also $f+g \in G$, und analog $\lambda f \in G$ für jedes $\lambda \in \mathbb{R}$. Auch die Menge der *ungeraden Funktionen*,

$$U := \{f: X \rightarrow \mathbb{R} \mid \forall x \in \mathbb{R} : f(-x) = -f(x)\},$$

bildet einen Teilraum von $F(X, \mathbb{R})$. Für den Durchschnitt gilt $G \cap U = \{0\}$, d.h. die einzige zugleich gerade und ungerade Funktion ist die konstante Nullfunktion. Ist nämlich $f \in G \cap U$ und $x \in X$, dann folgt $f(x) = f(-x) = -f(x)$, also $2f(x) = 0$ und somit $f(x) = 0$, d.h. $f = 0$.

II.2.15. BEISPIEL. Für jede Menge X , ist $\{f: X \rightarrow \mathbb{R} \mid \forall x \in X : f(x) \geq 0\}$ zwar abgeschlossen unter Addition, bildet jedoch keinen Teilraum von $F(X, \mathbb{R})$. Dasselbe gilt für die Teilmenge $F(X, \mathbb{Z}) \subseteq F(X, \mathbb{R})$, d.h. die Menge der Funktionen mit ganzzahligen Werten. Auch ist $F(X; \mathbb{R})$ kein Teilraum von $F(X; \mathbb{C})$, denn für $0 \neq f \in F(X; \mathbb{R})$ gilt $if \notin F(X; \mathbb{R})$.

II.3. Lineare Abbildungen. Lineare Abbildungen sind Abbildungen zwischen Vektorräumen, die mit der Vektorraumstruktur, d.h. Addition und Skalarmultiplikation, verträglich sind.

II.3.1. DEFINITION (Lineare Abbildungen). Eine Abbildung zwischen \mathbb{K} -Vektorräumen, $\varphi: V \rightarrow W$, wird *linear* genannt, wenn sie die folgenden beiden Eigenschaften besitzt:

- (a) $\forall v_1, v_2 \in V : \varphi(v_1 + v_2) = \varphi(v_1) + \varphi(v_2)$
- (b) $\forall \lambda \in \mathbb{K} \forall v \in V : \varphi(\lambda v) = \lambda \varphi(v)$

Die Menge aller linearen Abbildungen von V nach W wird mit $L(V, W)$ bezeichnet. Eine lineare Abbildung $V \rightarrow V$ wird auch *Endomorphismus* genannt. Die Menge aller Endomorphismen von V werden wir mit $\text{end}(V)$ bezeichnen.

II.3.2. BEMERKUNG. Für jede lineare Abbildung $\varphi: V \rightarrow W$ gilt $\varphi(0) = 0$. Aus der Linearität erhalten wir nämlich $\varphi(0) = \varphi(0 + 0) = \varphi(0) + \varphi(0)$ und Addition von $-\varphi(0)$ auf beiden Seiten liefert die gewünschte Gleichung, $0 = \varphi(0)$.

II.3.3. PROPOSITION (Komposition linearer Abbildungen). *Sind V, W und U drei \mathbb{K} -Vektorräume dann gilt:*

- (a) *Die identische Abbildung $\text{id}_V: V \rightarrow V$, $\text{id}_V(v) := v$, ist linear.*
- (b) *Für je zwei lineare Abbildungen $\varphi: V \rightarrow W$ und $\psi: W \rightarrow U$ ist auch die Komposition $\psi \circ \varphi: V \rightarrow U$ linear.*

(c) Ist $\varphi : V \rightarrow W$ eine bijektive lineare Abbildung, dann ist auch die Umkehrabbildung $\varphi^{-1} : W \rightarrow V$ linear.

BEWEIS. Behauptung (a) ist trivial. Ad (b): Sind $v_1, v_2 \in V$ so gilt:

$$\begin{aligned} (\psi \circ \varphi)(v_1 + v_2) &= \psi(\varphi(v_1 + v_2)) = \psi(\varphi(v_1) + \varphi(v_2)) \\ &= \psi(\varphi(v_1)) + \psi(\varphi(v_2)) = (\psi \circ \varphi)(v_1) + (\psi \circ \varphi)(v_2) \end{aligned}$$

Für $\lambda \in \mathbb{K}$ und $v \in V$ erhalten wir analog $(\psi \circ \varphi)(\lambda v) = \psi(\varphi(\lambda v)) = \psi(\lambda \varphi(v)) = \lambda \psi(\varphi(v)) = \lambda(\psi \circ \varphi)(v)$. Dies zeigt, dass die Komposition $\psi \circ \varphi$ linear ist. Um (c) zu verifizieren sei nun $\varphi : V \rightarrow W$ eine lineare Bijektion mit Umkehrabbildung $\varphi^{-1} : W \rightarrow V$. Sind $w_1, w_2 \in W$ so gilt:

$$\begin{aligned} \varphi^{-1}(w_1 + w_2) &= \varphi^{-1}(\varphi(\varphi^{-1}(w_1)) + \varphi(\varphi^{-1}(w_2))) && \text{denn } \forall w \in W : w = \varphi(\varphi^{-1}(w)) \\ &= \varphi^{-1}(\varphi(\varphi^{-1}(w_1) + \varphi^{-1}(w_2))) && \text{Linearität von } \varphi \\ &= \varphi^{-1}(w_1) + \varphi^{-1}(w_2) && \text{denn } \forall v \in V : \varphi^{-1}(\varphi(v)) = v \end{aligned}$$

Für $\lambda \in \mathbb{K}$ und $w \in W$ erhalten wir analog $\varphi^{-1}(\lambda w) = \varphi^{-1}(\lambda \varphi(\varphi^{-1}(w))) = \varphi^{-1}(\varphi(\lambda \varphi^{-1}(w))) = \lambda \varphi^{-1}(w)$. Dies zeigt, dass die Umkehrabbildung φ^{-1} linear ist. \square

II.3.4. BEISPIEL. Sind $n, m \in \mathbb{N}$ und $a_{ij} \in \mathbb{K}$, dann ist die durch

$$\psi \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} := \begin{pmatrix} a_{11}x_1 + \cdots + a_{1n}x_n \\ \vdots \\ a_{m1}x_1 + \cdots + a_{mn}x_n \end{pmatrix} = x_1 \begin{pmatrix} a_{11} \\ \vdots \\ a_{m1} \end{pmatrix} + \cdots + x_n \begin{pmatrix} a_{1n} \\ \vdots \\ a_{mn} \end{pmatrix}$$

definierte Abbildung $\psi : \mathbb{K}^n \rightarrow \mathbb{K}^m$ linear, vgl. Übungsaufgabe 1. Wir werden später sehen, dass jede lineare Abbildung $\mathbb{K}^n \rightarrow \mathbb{K}^m$ von dieser Form ist, vgl. Satz II.4.4 unten.

II.3.5. BEISPIEL (Inklusionen). Für jeden Teilraum W eines Vektorraums V ist die *kanonische Inklusionsabbildung*, $W \rightarrow V$, $w \mapsto w$, offensichtlich linear. Die Verknüpfungen auf W wurden genau so definiert, dass diese Inklusionsabbildung linear wird.

II.3.6. BEISPIEL (Koordinatenprojektionen). Für jedes $i = 1, \dots, n$ ist die Abbildung $p_i : \mathbb{K}^n \rightarrow \mathbb{K}$, $p_i(x) := x_i$, die einem Vektor seine i -te Koordinate zuordnet, linear. Sie wird als i -te *Koordinatenprojektion* bezeichnet.

II.3.7. BEISPIEL (Polynome als Funktionen). Ist $p = p_0 + p_1z + p_2z^2 + \cdots$ ein Polynom in $\mathbb{K}[z]$ und $x \in \mathbb{K}$, so können wir x in p einsetzen und erhalten eine Zahl $p(x) \in \mathbb{K}$,

$$p(x) := p_0 + p_1x + p_2x^2 + \cdots$$

wobei die rechte Seite eine endliche Summe in \mathbb{K} darstellt. Jedes Polynom $p \in \mathbb{K}[z]$ liefert daher eine Funktion $\mathbb{K} \rightarrow \mathbb{K}$, $x \mapsto p(x)$. Wir erhalten somit eine Abbildung

$\phi: \mathbb{K}[z] \rightarrow F(\mathbb{K}, \mathbb{K})$, die einem Polynom $p \in \mathbb{K}[z]$ die Funktion $x \mapsto p(x)$ zuordnet. Diese Abbildung ist linear d.h. es gilt $\phi(p+q) = \phi(p) + \phi(q)$ sowie $\phi(\lambda p) = \lambda\phi(p)$, für beliebige $p, q \in \mathbb{K}[z]$ und $\lambda \in \mathbb{K}$. In anderen Worten, für jedes $x \in \mathbb{K}$ ist

$$(p+q)(x) = p(x) + q(x) \quad \text{und} \quad (\lambda p)(x) = \lambda p(x).$$

Darüber hinaus gilt auch $\phi(pq) = \phi(p)\phi(q)$, d.h. für jedes $x \in \mathbb{K}$ haben wir

$$(pq)(x) = p(x)q(x).$$

Die Abbildung $\phi: \mathbb{K}[z] \rightarrow F(\mathbb{K}, \mathbb{K})$ ist daher sogar ein Homomorphismus von Algebren, d.h. eine lineare Abbildung die auch mit der Multiplikation verträglich ist, vgl. Beispiel II.1.7 und Beispiel II.1.9. Übungsaufgabe 42 behandelt eine Verallgemeinerung dieses Beispiels.

II.3.8. BEISPIEL. Sei V ein Vektorraum und X eine Menge. Für jedes $x \in X$ ist die Abbildung $ev_x: F(X, V) \rightarrow V$, $ev_x(f) := f(x)$, linear. Für jede Teilmenge $A \subseteq X$ ist auch die Abbildung $F(X, V) \rightarrow F(A, V)$, $f \mapsto f|_A$, linear. Allgemeiner ist für jede Abbildung $g: Y \rightarrow X$ die Zuordnung $F(X, V) \rightarrow F(Y, V)$, $f \mapsto f \circ g$, eine lineare Abbildung, vgl. Übungsaufgabe 24.

II.3.9. BEISPIEL (Lineare Abbildungen aus der Analysis). Es seien $a < b$ zwei reelle Zahlen. Für jedes $k \in \mathbb{N}$ liefert die Ableitung eine lineare Abbildung,

$$C^k((a, b), \mathbb{R}) \rightarrow C^{k-1}((a, b), \mathbb{R}), \quad f \mapsto f',$$

denn $(f+g)' = f' + g'$ und $(\lambda f)' = \lambda f'$ für alle $f, g \in C^k((a, b), \mathbb{R})$ und $\lambda \in \mathbb{R}$. Auch das Integral liefert eine lineare Abbildung,

$$C([a, b], \mathbb{R}) \rightarrow \mathbb{R}, \quad f \mapsto \int_a^b f(x) dx,$$

denn es gilt $\int_a^b (f+g)(x) dx = \int_a^b f(x) dx + \int_a^b g(x) dx$ und $\int_a^b (\lambda f)(x) dx = \lambda \int_a^b f(x) dx$ für alle $f, g \in C([a, b], \mathbb{R})$ und $\lambda \in \mathbb{R}$. Bezeichnet $F_{\text{conv}}(\mathbb{N}, \mathbb{R}) \subseteq F(\mathbb{N}, \mathbb{R})$ den Teilraum der konvergenten Folgen, dann ist die Abbildung

$$F_{\text{conv}}(\mathbb{N}, \mathbb{R}) \rightarrow \mathbb{R}, \quad (x_n)_{n \in \mathbb{N}} \mapsto \lim_{n \rightarrow \infty} x_n,$$

linear, denn es gilt $\lim_{n \rightarrow \infty} (x_n + y_n) = \lim_{n \rightarrow \infty} x_n + \lim_{n \rightarrow \infty} y_n$ und $\lim_{n \rightarrow \infty} \lambda x_n = \lambda \lim_{n \rightarrow \infty} x_n$ für je zwei konvergente Folgen $(x_n)_{n \in \mathbb{N}}$, $(y_n)_{n \in \mathbb{N}}$ und $\lambda \in \mathbb{R}$.

II.3.10. DEFINITION (Isomorphismen). Eine Bijektion zwischen \mathbb{K} -Vektorräumen, $\varphi: V \rightarrow W$, wird (*linearer*) *Isomorphismus* genannt, wenn $\varphi: V \rightarrow W$ und ihre Umkehrabbildung $\varphi^{-1}: W \rightarrow V$ beide linear sind. Existiert ein Isomorphismus zwischen V und W , dann werden V und W *isomorph* genannt und wir schreiben $V \cong W$.

II.3.11. BEMERKUNG. Nach Proposition II.3.3(c) ist jede lineare Bijektion schon ein Isomorphismus, die Umkehrabbildung ist dann automatisch linear.

Sind zwei Vektorräume isomorph, $V \cong W$, dann können sie vom Standpunkt der linearen Algebra als im Wesentlichen gleich betrachtet werden. Bis auf Umbenennung der Elemente mit Hilfe eines Isomorphismus $V \cong W$, entsprechen die Vektorraumoperationen in V ja genau den Vektorraumoperationen in W .

Für jeden Vektorraum V ist die identische Abbildung, $\text{id}_V: V \rightarrow V$, ein Isomorphismus, $\text{id}_V^{-1} = \text{id}_V$, es gilt daher $V \cong V$. Ist $\varphi: V \rightarrow W$ ein Isomorphismus, dann ist auch die Umkehrabbildung $\varphi^{-1}: W \rightarrow V$ ein Isomorphismus, $(\varphi^{-1})^{-1} = \varphi$, aus $V \cong W$ folgt daher stets $W \cong V$. Sind $\varphi: V \rightarrow W$ und $\psi: W \rightarrow U$ zwei Isomorphismen, dann ist auch deren Komposition $\psi \circ \varphi: V \rightarrow U$ ein Isomorphismus. Die letzte Behauptung folgt aus Proposition II.3.3(b) und der Formel für die Umkehrabbildung einer Komposition, $(\psi \circ \varphi)^{-1} = \varphi^{-1} \circ \psi^{-1}$. Mit $V \cong W$ und $W \cong U$ gilt daher auch $V \cong U$.

Die Menge aller invertierbaren linearen Abbildungen $V \rightarrow V$ wird mit $\text{GL}(V)$ bezeichnet. Aus dem vorangehenden Absatz folgt sofort, dass $\text{GL}(V)$ bezüglich der Komposition von Abbildungen eine Gruppe bildet. Sie wird die *allgemeine lineare Gruppe* genannt. Diese Gruppe ist i.A. nicht abelsch, d.h. für $\varphi, \psi \in \text{GL}(V)$ gilt i.A. $\varphi \circ \psi \neq \psi \circ \varphi$, vgl. Übungsaufgabe 25.

II.3.12. BEISPIEL. Betrachte den Teilraum

$$W := \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{K}^3 : x + 2y + 3z = 0 \right\}$$

von \mathbb{K}^3 . Dann ist die Abbildung

$$\mathbb{K}^2 \xrightarrow{\cong} W, \quad \begin{pmatrix} \lambda \\ \mu \end{pmatrix} \mapsto \lambda \begin{pmatrix} -2 \\ 1 \\ 0 \end{pmatrix} + \mu \begin{pmatrix} -3 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} -2\lambda - 3\mu \\ \lambda \\ \mu \end{pmatrix}$$

eine lineare Bijektion, also ein Isomorphismus. Es gilt daher $W \cong \mathbb{K}^2$. Wir werden später sehen, dass jeder Teilraum von \mathbb{K}^n isomorph zu \mathbb{K}^m ist, für ein eindeutig bestimmtes $m \in \mathbb{N}_0$, für das auch $0 \leq m \leq n$ gelten muss.

II.3.13. BEISPIEL. Sind $m < n$ zwei natürliche Zahlen, dann bildet

$$W := \left\{ \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{K}^n \mid \begin{array}{l} x_{m+1} = 0 \\ \vdots \\ x_n = 0 \end{array} \right\}$$

einen Teilraum von \mathbb{K}^n , der im Wesentlichen mit \mathbb{K}^m übereinstimmt. Genauer ist $\varphi: \mathbb{K}^m \rightarrow W$, $\varphi(x) := \begin{pmatrix} x \\ 0 \end{pmatrix}$, eine lineare Bijektion, also ein Isomorphismus.

II.3.14. BEISPIEL. Für $0 \neq x \in \mathbb{K}^n$ bildet $W := \{\lambda x : \lambda \in \mathbb{K}\}$ einen Teilraum von \mathbb{K}^n , der zu \mathbb{K} isomorph ist. Die Abbildung $\varphi: \mathbb{K} \rightarrow W$, $\lambda \mapsto \lambda x$, ist eine lineare Bijektion, also ein Isomorphismus, vgl. Bemerkung II.3.11.

II.3.15. BEISPIEL. Vektoren aus \mathbb{K}^n können mit Funktionen $\{1, \dots, n\} \rightarrow \mathbb{K}$ identifiziert werden, indem die i -te Komponente eines Vektors als Funktionswert bei i interpretiert wird. Dabei entspricht die Komponentenweise Addition von Vektoren in \mathbb{K}^n genau der punktweisen Addition von Funktionen, und analog für

die Skalarmultiplikation. Dies lässt sich wie folgt präzisieren. Die offensichtlich bijektive Abbildung,

$$\phi : F(\{1, \dots, n\}, \mathbb{K}) \xrightarrow{\cong} \mathbb{K}^n, \quad \phi(f) := \begin{pmatrix} f(1) \\ \vdots \\ f(n) \end{pmatrix},$$

ist ein linearer Isomorphismus, denn für $f, g \in F(\{1, \dots, n\}, \mathbb{K})$ gilt

$$\phi(f+g) = \begin{pmatrix} (f+g)(1) \\ \vdots \\ (f+g)(n) \end{pmatrix} = \begin{pmatrix} f(1)+g(1) \\ \vdots \\ f(n)+g(n) \end{pmatrix} = \begin{pmatrix} f(1) \\ \vdots \\ f(n) \end{pmatrix} + \begin{pmatrix} g(1) \\ \vdots \\ g(n) \end{pmatrix} = \phi(f) + \phi(g)$$

und analog auch $\phi(\lambda f) = \lambda \phi(f)$, für alle $\lambda \in \mathbb{K}$. Nach Bemerkung II.3.11 ist die Umkehrabbildung automatisch linear.

II.3.16. BEISPIEL. Es bezeichne $F_0(\mathbb{N}_0; \mathbb{K}) \subseteq F(\mathbb{N}_0, \mathbb{K})$ den Teilraum aller Folgen, deren Folgenglieder fast alle gleich 0 sind. Ordnen wir einem Polynom die Folge seiner Koeffizienten zu, so erhalten wir einen linearen Isomorphismus

$$\mathbb{K}[z] \cong F_0(\mathbb{N}_0; \mathbb{K}), \quad p = p_0 + p_1 z + p_2 z^2 + \dots \leftrightarrow (p_i)_{i \in \mathbb{N}_0}$$

Analog erhalten wir einen Isomorphismus $\mathbb{K}[z]_{\leq n} \cong \mathbb{K}^{n+1}$, wobei $\mathbb{K}[z]_{\leq n}$ den Teilraum aller Polynome vom Grad kleiner oder gleich n bezeichnet.

II.3.17. BEISPIEL. Die Vektorräume $\mathbb{K} = \mathbb{K}^1$ und \mathbb{K}^2 sind nicht isomorph, denn eine lineare Abbildung $\mathbb{K} \rightarrow \mathbb{K}^2$ kann niemals surjektiv sein. Um dies einzusehen, sei $\varphi : \mathbb{K} \rightarrow \mathbb{K}^2$ linear. Wir bezeichnen die beiden Komponenten von $\varphi(1)$ mit x_1 und x_2 , d.h. $\varphi(1) = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$. Für jedes $\lambda \in \mathbb{K}$ gilt dann wegen der Linearität $\varphi(\lambda) = \lambda \varphi(1) = \lambda \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} \lambda x_1 \\ \lambda x_2 \end{pmatrix}$. Liegt der Vektor $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$ im Bild von φ , dann existiert $\lambda \in \mathbb{K}$ mit $\varphi(\lambda) = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ und aus der Formel oben folgt $\begin{pmatrix} \lambda x_1 \\ \lambda x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$, somit $x_1 \neq 0$ und $x_2 = 0$. Liegt $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$ im Bild von φ , dann folgt analog $x_1 = 0$ und $x_2 \neq 0$. Liegen beide Vektoren, $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ und $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$, im Bild von φ erhalten wir einen Widerspruch. Die Abbildung φ kann daher nicht surjektiv sein. Damit ist $\mathbb{K} \not\cong \mathbb{K}^2$ gezeigt. Wir werden später sehen, dass \mathbb{K}^n und \mathbb{K}^m genau dann isomorph sind, wenn $n = m$ gilt.

II.3.18. PROPOSITION. Sind V und W zwei Vektorräume über \mathbb{K} , dann ist die Menge aller linearen Abbildungen, $L(V, W)$, ein Teilraum von $F(V, W)$. Insbesondere bildet $L(V, W)$ bezüglich punktweiser Addition und Skalarmultiplikation,

$$(\varphi_1 + \varphi_2)(v) := \varphi_1(v) + \varphi_2(v) \quad \text{und} \quad (\lambda \varphi)(v) := \lambda \varphi(v)$$

einen \mathbb{K} -Vektorraum, $\varphi, \varphi_1, \varphi_2 \in L(V, W)$, $\lambda \in \mathbb{K}$, $v \in V$. Die Vektorraumstruktur ist in folgendem Sinn mit der Komposition linearer Abbildungen verträglich: Für beliebige $\varphi, \varphi_1, \varphi_2 \in L(V, W)$, $\psi, \psi_1, \psi_2 \in L(W, U)$ und $\lambda \in \mathbb{K}$ gilt:

- (a) $(\psi_1 + \psi_2) \circ \varphi = \psi_1 \circ \varphi + \psi_2 \circ \varphi$ und $(\lambda \psi) \circ \varphi = \lambda(\psi \circ \varphi)$, sowie
- (b) $\psi \circ (\varphi_1 + \varphi_2) = \psi \circ \varphi_1 + \psi \circ \varphi_2$ und $\psi \circ (\lambda \varphi) = \lambda(\psi \circ \varphi)$.

BEWEIS. Zunächst ist $L(V, W)$ nicht leer, denn die konstante Nullabbildung $0: V \rightarrow W$, $0(v) := 0$, ist offensichtlich linear. Um die Abgeschlossenheit unter Addition zu zeigen, seien nun $\varphi_1: V \rightarrow W$ und $\varphi_2: V \rightarrow W$ zwei lineare Abbildungen. Für beliebige $v_1, v_2 \in V$ gilt dann:

$$\begin{aligned} (\varphi_1 + \varphi_2)(v_1 + v_2) &= \varphi_1(v_1 + v_2) + \varphi_2(v_1 + v_2) \\ &= (\varphi_1(v_1) + \varphi_1(v_2)) + (\varphi_2(v_1) + \varphi_2(v_2)) \\ &= (\varphi_1(v_1) + \varphi_2(v_1)) + (\varphi_1(v_2) + \varphi_2(v_2)) \\ &= (\varphi_1 + \varphi_2)(v_1) + (\varphi_1 + \varphi_2)(v_2), \end{aligned}$$

Analog erhalten wir für jedes $\lambda \in \mathbb{K}$ und jedes $v \in V$ auch $(\varphi_1 + \varphi_2)(\lambda v) = \varphi_1(\lambda v) + \varphi_2(\lambda v) = \lambda\varphi_1(v) + \lambda\varphi_2(v) = \lambda(\varphi_1(v) + \varphi_2(v)) = \lambda(\varphi_1 + \varphi_2)(v)$. Dies zeigt, dass auch die Summe $\varphi_1 + \varphi_2$ eine lineare Abbildung ist, $L(V, W)$ ist daher abgeschlossen unter Addition. Für die Abgeschlossenheit unter Skalarmultiplikation sei nun $\varphi: V \rightarrow W$ linear und $\lambda \in \mathbb{K}$. Für beliebige $v_1, v_2 \in V$ gilt dann:

$$\begin{aligned} (\lambda\varphi)(v_1 + v_2) &= \lambda\varphi(v_1 + v_2) = \lambda(\varphi(v_1) + \varphi(v_2)) \\ &= \lambda\varphi(v_1) + \lambda\varphi(v_2) = (\lambda\varphi)(v_1) + (\lambda\varphi)(v_2). \end{aligned}$$

Für $\mu \in \mathbb{K}$ und $v \in V$ erhalten wir analog $(\lambda\varphi)(\mu v) = \lambda\varphi(\mu v) = \lambda(\mu\varphi(v)) = (\lambda\mu)\varphi(v) = (\mu\lambda)\varphi(v) = \mu(\lambda\varphi(v)) = \mu(\lambda\varphi)(v)$. Somit ist $\lambda\varphi$ linear, und $L(V, W)$ daher abgeschlossen unter Skalarmultiplikation. Dies zeigt, dass $L(V, W)$ einen Teilraum von $F(V, W)$ bildet.

Ad (a): Für $\varphi \in L(V, W)$, $\psi_1, \psi_2 \in L(W, U)$ und $v \in V$ erhalten wir

$$\begin{aligned} ((\psi_1 + \psi_2) \circ \varphi)(v) &= (\psi_1 + \psi_2)(\varphi(v)) = \psi_1(\varphi(v)) + \psi_2(\varphi(v)) \\ &= (\psi_1 \circ \varphi)(v) + (\psi_2 \circ \varphi)(v) = (\psi_1 \circ \varphi + \psi_2 \circ \varphi)(v). \end{aligned}$$

Da dies für beliebige $v \in V$ gilt, folgt $(\psi_1 + \psi_2) \circ \varphi = \psi_1 \circ \varphi + \psi_2 \circ \varphi$. Analog ist $((\lambda\psi) \circ \varphi)(v) = (\lambda\psi)(\varphi(v)) = \lambda\psi(\varphi(v)) = \lambda(\psi \circ \varphi)(v) = (\lambda(\psi \circ \varphi))(v)$, also $(\lambda\psi) \circ \varphi = \lambda(\psi \circ \varphi)$, für alle $\varphi \in L(V, W)$, $\psi \in L(W, U)$ und $\lambda \in \mathbb{K}$.

Ad (b): Für $\varphi_1, \varphi_2 \in L(V, W)$, $\psi \in L(W, U)$ und $v \in V$ erhalten wir

$$\begin{aligned} (\psi \circ (\varphi_1 + \varphi_2))(v) &= \psi((\varphi_1 + \varphi_2)(v)) \\ &= \psi(\varphi_1(v) + \varphi_2(v)) \\ &= \psi(\varphi_1(v)) + \psi(\varphi_2(v)) \\ &= (\psi \circ \varphi_1)(v) + (\psi \circ \varphi_2)(v) \\ &= (\psi \circ \varphi_1 + \psi \circ \varphi_2)(v), \end{aligned}$$

also $\psi \circ (\varphi_1 + \varphi_2) = \psi \circ \varphi_1 + \psi \circ \varphi_2$. Analog gilt $(\psi \circ (\lambda\varphi))(v) = \psi((\lambda\varphi)(v)) = \psi(\lambda\varphi(v)) = \lambda\psi(\varphi(v)) = \lambda(\psi \circ \varphi)(v) = (\lambda(\psi \circ \varphi))(v)$, also $\psi \circ (\lambda\varphi) = \lambda(\psi \circ \varphi)$, für alle $\varphi \in L(V, W)$, $\psi \in L(W, U)$ und $\lambda \in \mathbb{K}$. \square

II.3.19. BEMERKUNG. Nach Proposition II.3.18 oben bildet $\text{end}(V)$ eine assoziative \mathbb{K} -Algebra mit Eins. Diese Algebra ist i.A. nicht kommutativ. Betrachten wir etwa die beiden linearen Abbildungen $\varphi: \mathbb{K}^2 \rightarrow \mathbb{K}^2$, $\varphi \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} := \begin{pmatrix} x_2 \\ 0 \end{pmatrix}$, und $\psi: \mathbb{K}^2 \rightarrow \mathbb{K}^2$, $\psi \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} := \begin{pmatrix} 0 \\ x_1 \end{pmatrix}$, dann gilt $(\varphi \circ \psi) \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} x_1 \\ 0 \end{pmatrix}$, sowie $(\psi \circ \varphi) \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ x_2 \end{pmatrix}$, also $(\varphi \circ \psi) \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \neq \begin{pmatrix} 0 \\ 0 \end{pmatrix} = (\psi \circ \varphi) \begin{pmatrix} 1 \\ 0 \end{pmatrix}$, und daher $\varphi \circ \psi \neq \psi \circ \varphi$.

II.3.20. PROPOSITION. *Ist $\varphi: V \rightarrow W$ eine lineare Abbildung, dann gilt:*

- (a) *Für jeden Teilraum W' von W ist auch $\varphi^{-1}(W')$ Teilraum von V .*
 (b) *Für jeden Teilraum V' von V ist auch $\varphi(V')$ Teilraum von W .*

BEWEIS. Ad (a): Zunächst ist $\varphi^{-1}(W')$ nicht leer, denn aus $\varphi(0) = 0 \in W'$ folgt $0 \in \varphi^{-1}(W')$. Für $v_1, v_2 \in \varphi^{-1}(W')$ folgt $\varphi(v_1), \varphi(v_2) \in W'$, also $\varphi(v_1 + v_2) = \varphi(v_1) + \varphi(v_2) \in W'$ und somit $v_1 + v_2 \in \varphi^{-1}(W')$. Dies zeigt, dass $\varphi^{-1}(W')$ abgeschlossen unter Addition ist. Sind nun $\lambda \in \mathbb{K}$ und $v \in \varphi^{-1}(W')$, dann folgt $\varphi(v) \in W'$, also $\varphi(\lambda v) = \lambda\varphi(v) \in W'$ und daher $\lambda v \in \varphi^{-1}(W')$. Dies zeigt, dass $\varphi^{-1}(W')$ auch abgeschlossen unter Skalarmultiplikation ist.

Ad (b): Zunächst ist $\varphi(V')$ nicht leer, denn $0 = \varphi(0) \in \varphi(V')$. Seien nun $w_1, w_2 \in \varphi(V')$. Es existieren daher $v_1, v_2 \in V'$ mit $\varphi(v_1) = w_1$ und $\varphi(v_2) = w_2$. Wir erhalten $w_1 + w_2 = \varphi(v_1) + \varphi(v_2) = \varphi(v_1 + v_2) \in \varphi(V')$, denn $v_1 + v_2 \in V'$. Dies zeigt, dass $\varphi(V')$ abgeschlossen unter Addition ist. Ist $\lambda \in \mathbb{K}$ und $w \in \varphi(V')$, dann existiert $v \in V'$ mit $\varphi(v) = w$ und wir erhalten $\lambda w = \lambda\varphi(v) = \varphi(\lambda v) \in \varphi(V')$, denn $\lambda v \in V'$. Somit ist $\varphi(V')$ auch abgeschlossen unter Skalarmultiplikation. \square

II.3.21. DEFINITION (Kern und Bild). Unter dem *Kern* einer linearen Abbildung $\varphi: V \rightarrow W$ verstehen wir den Teilraum

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

Unter dem *Bild* von φ verstehen wir den Teilraum

$$\text{img}(\varphi) := \varphi(V) = \{\varphi(v) \mid v \in V\} \subseteq W.$$

Nach Proposition II.3.20 sind beides tatsächlich Teilräume.

Wir beenden diesen Abschnitt mit folgendem einfachen Resultat, das oft verwendet wird um die Injektivität linearer Abbildungen zu überprüfen.

II.3.22. PROPOSITION. *Eine lineare Abbildung $\varphi: V \rightarrow W$ ist genau dann injektiv wenn sie trivialen Kern hat, d.h. genau dann wenn $\ker(\varphi) = \{0\}$ gilt.*

BEWEIS. Ist φ injektiv, dann besteht $\ker(\varphi) = \varphi^{-1}(0)$ aus höchstens einem Element, und da $\varphi(0) = 0$, folgt $\ker(\varphi) = \{0\}$. Dies zeigt die eine Implikation. Sei nun umgekehrt $\ker(\varphi) = \{0\}$. Um zu zeigen, dass φ injektiv ist betrachten wir $v_1, v_2 \in V$ mit $\varphi(v_1) = \varphi(v_2)$. Aus der Linearität von φ erhalten wir $\varphi(v_2 - v_1) = \varphi(v_2) - \varphi(v_1) = 0$, also $v_2 - v_1 \in \ker(\varphi)$, somit $v_2 - v_1 = 0$ und daher $v_1 = v_2$. Damit ist auch die umgekehrte Implikation gezeigt. \square

II.4. Matrizen. Sei \mathbb{K} ein Körper. Wir wollen in diesem Abschnitt lineare Abbildungen $\mathbb{K}^n \rightarrow \mathbb{K}^m$ mit Hilfe von Matrizen beschreiben. Unter einer $(m \times n)$ -*Matrix* über einem Körper \mathbb{K} verstehen wir ein rechteckiges Schema der Form

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

mit Eintragungen $a_{ij} \in \mathbb{K}$. Die Menge aller $(m \times n)$ -Matrizen über \mathbb{K} bezeichnen wir mit $M_{m \times n}(\mathbb{K})$. Die Summe zweier $(m \times n)$ -Matrizen wird elementweise, d.h. durch

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

definiert. Beachte, dass die Summe zweier Matrizen nur dann definiert ist, wenn sie gleiche Zeilen- und Spaltenzahl haben. Auch die Skalarmultiplikation mit $\lambda \in \mathbb{K}$ ist elementweise definiert, d.h.

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

Mit diesen Operationen wird $M_{m \times n}(\mathbb{K})$ zu einem Vektorraum über \mathbb{K} . Fassen die Eintragungen einer $(m \times n)$ -Matrix in einem Spaltenvektor mit mn vielen Komponenten zusammen, so erhalten wir einen Isomorphismus von \mathbb{K} -Vektorräumen,

$$M_{m \times n}(\mathbb{K}) \cong \mathbb{K}^{mn}.$$

II.4.1. BEISPIEL. Etwa ist

$$\begin{pmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{pmatrix} + \begin{pmatrix} 0 & 2 \\ 1 & 3 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 4 \\ 4 & 7 \\ 5 & 7 \end{pmatrix} \quad \text{und} \quad \frac{1}{5} \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix} = \begin{pmatrix} 1/5 & 2/5 & 3/5 \\ 4/5 & 1 & 6/5 \\ 7/5 & 8/5 & 9/5 \end{pmatrix}.$$

Unter dem *Produkt* einer $(m \times n)$ -Matrix mit einer $(n \times l)$ -Matrix verstehen wir die $(m \times l)$ -Matrix

$$\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} b_{11} & \cdots & b_{1l} \\ \vdots & & \vdots \\ b_{n1} & \cdots & b_{nl} \end{pmatrix} := \begin{pmatrix} \sum_{k=1}^n a_{1k} b_{k1} & \cdots & \sum_{k=1}^n a_{1k} b_{kl} \\ \vdots & & \vdots \\ \sum_{k=1}^n a_{mk} b_{k1} & \cdots & \sum_{k=1}^n a_{mk} b_{kl} \end{pmatrix}$$

Beachte, dass das Matrizenprodukt nur definiert ist, wenn die Spaltenzahl der linken Matrix mit der Zeilenzahl der rechten Matrix übereinstimmt. Die Bedeutung dieses Produkts wird in Satz II.4.4 unten klar werden. Unter der $(n \times n)$ -Einheitsmatrix, $I_n \in M_{n \times n}(\mathbb{K})$, verstehen wir die Matrix

$$I_n := \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix},$$

wobei die Diagonaleinträge gleich 1 und alle anderen Eintragungen gleich 0 sind. Die Einheitsmatrizen sind neutrale Element der Matrizenmultiplikation.

II.4.2. BEISPIEL. Etwa ist

$$\begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 4 \\ 3 & 4 & 5 \\ 4 & 5 & 6 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 5 & 3 \\ 7 & 5 \\ 9 & 7 \\ 11 & 9 \end{pmatrix} \quad \text{und} \quad \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix} \begin{pmatrix} 2 \\ -1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 3 \end{pmatrix}.$$

Um eine kompaktere Schreibweise zur Verfügung zu haben, bezeichnen wir den Eintrag in der i -ten Zeile und j -ten Spalte einer Matrix A mit A_{ij} . Für $A \in M_{m \times n}(\mathbb{K})$ und $B \in M_{n \times l}(\mathbb{K})$ gilt daher nach Definition des Matrizenprodukts

$$(AB)_{ij} = \sum_{k=1}^n A_{ik}B_{kj}, \quad 1 \leq i \leq m, 1 \leq j \leq l.$$

Addition und Skalarmultiplikation lassen sich damit wie folgt schreiben,

$$(A + A')_{ij} = A_{ij} + A'_{ij} \quad (\lambda A)_{ij} = \lambda A_{ij}, \quad 1 \leq i \leq m, 1 \leq j \leq n,$$

wobei $A, A' \in M_{m \times n}(\mathbb{K})$ und $\lambda \in \mathbb{K}$. Für die Einheitsmatrix erhalten wir

$$(I_n)_{ij} = \delta_{ij} = \begin{cases} 1 & \text{falls } i = j, \text{ und} \\ 0 & \text{falls } i \neq j. \end{cases} \quad (\text{II.1})$$

Das Symbol δ_{ij} wird *Kronecker-Symbol* genannt.

II.4.3. PROPOSITION (Rechenregeln für Matrizen). *Sei \mathbb{K} ein Körper. Für Matrizen $A, A' \in M_{m \times n}(\mathbb{K})$, $B, B' \in M_{n \times l}(\mathbb{K})$, $C \in M_{l \times k}(\mathbb{K})$ und $\lambda \in \mathbb{K}$ gilt:*

- (a) $(AB)C = A(BC)$
- (b) $I_m A = A = A I_n$
- (c) $(A + A')B = AB + A'B$ und $(\lambda A)B = \lambda(AB)$
- (d) $A(B + B') = AB + AB'$ und $A(\lambda B) = \lambda(AB)$

BEWEIS. Ad (a): Für $1 \leq i \leq m$ und $1 \leq j \leq k$ gilt

$$\begin{aligned} ((AB)C)_{ij} &= \sum_{s=1}^l (AB)_{is} C_{sj} = \sum_{s=1}^l \left(\sum_{t=1}^n A_{it} B_{ts} \right) C_{sj} = \sum_{s=1}^l \sum_{t=1}^n A_{it} B_{ts} C_{sj} \\ &= \sum_{t=1}^n \sum_{s=1}^l A_{it} B_{ts} C_{sj} = \sum_{t=1}^n A_{it} \sum_{s=1}^l B_{ts} C_{sj} = \sum_{t=1}^n A_{it} (BC)_{tj} = (A(BC))_{ij} \end{aligned}$$

also $(AB)C = A(BC)$. Ad (b): Für $1 \leq i \leq m$ und $1 \leq j \leq n$ gilt wegen (II.1)

$$(I_m A)_{ij} = \sum_{k=1}^n (I_m)_{ik} A_{kj} = A_{ij},$$

also $I_m A = A$. Analog lässt sich $AI_n = A$ zeigen. Ad (c): Für $1 \leq i \leq m$ und $1 \leq j \leq l$ gilt

$$\begin{aligned} ((A + A')B)_{ij} &= \sum_{k=1}^n (A + A')_{ik} B_{kj} = \sum_{k=1}^n (A_{ik} + A'_{ik}) B_{kj} \\ &= \sum_{k=1}^n A_{ik} B_{kj} + \sum_{k=1}^n A'_{ik} B_{kj} = (AB)_{ij} + (A'B)_{ij} = (AB + A'B)_{ij} \end{aligned}$$

also $(A + A')B = AB + A'B$. Analog haben wir $((\lambda A)B)_{ij} = \sum_{k=1}^n (\lambda A)_{ik} B_{kj} = \sum_{k=1}^n \lambda A_{ik} B_{kj} = \lambda \sum_{k=1}^n A_{ik} B_{kj} = \lambda (AB)_{ij} = (\lambda(AB))_{ij}$, also $(\lambda A)B = \lambda(AB)$. Die letzte Behauptung (d) lässt sich analog beweisen. \square

II.4.4. SATZ. *Es sei \mathbb{K} ein Körper und $n, m \in \mathbb{N}$. Jede Matrix $A \in M_{m \times n}(\mathbb{K})$ definiert eine lineare Abbildung $\psi_A: \mathbb{K}^n \rightarrow \mathbb{K}^m$, $\psi_A(x) := Ax$, und jede lineare Abbildung $\mathbb{K}^n \rightarrow \mathbb{K}^m$ ist von dieser Form für eine eindeutig bestimmte Matrix $A \in M_{m \times n}(\mathbb{K})$. Dabei stimmt der i -te Spaltenvektor von A mit dem Bild des i -ten Einheitsvektors, $\psi_A(e_i) \in \mathbb{K}^m$, überein. Diese Zuordnung,*

$$M_{m \times n}(\mathbb{K}) \cong L(\mathbb{K}^n, \mathbb{K}^m), \quad A \leftrightarrow \psi_A,$$

ist ein linearer Isomorphismus, es gilt daher

$$\psi_{A+A'} = \psi_A + \psi_{A'} \quad \text{und} \quad \psi_{\lambda A} = \lambda \psi_A,$$

für beliebige $A, A' \in M_{m \times n}(\mathbb{K})$ und $\lambda \in \mathbb{K}$. Weiters haben wir

$$\psi_{AB} = \psi_A \circ \psi_B \quad \text{sowie} \quad \psi_{I_n} = \text{id}_{\mathbb{K}^n},$$

für alle $A \in M_{m \times n}(\mathbb{K})$ und $B \in M_{n \times k}(\mathbb{K})$.

BEWEIS. Die Abbildung ψ_A ist linear, denn nach Proposition II.4.3(d) gilt $\psi_A(x + y) = A(x + y) = Ax + Ay = \psi_A(x) + \psi_A(y)$ und $\psi_A(\lambda x) = A(\lambda x) = \lambda Ax = \lambda \psi_A(x)$, wobei $x, y \in \mathbb{K}^n$, $\lambda \in \mathbb{K}$. Offensichtlich ist $\psi_A(e_i) = Ae_i$ gerade der i -te Spaltenvektor von A . Die Matrix A ist also durch die lineare Abbildung ψ_A eindeutig bestimmt, die Zuordnung $M_{m \times n}(\mathbb{K}) \rightarrow L(\mathbb{K}^n, \mathbb{K}^m)$, $A \mapsto \psi_A$ ist daher injektiv.

Wir werden nun zeigen, dass diese Zuordnung auch surjektiv ist. Sei dazu $\varphi: \mathbb{K}^n \rightarrow \mathbb{K}^m$ eine beliebige lineare Abbildung. Es bezeichne $A \in M_{m \times n}(\mathbb{K})$ jene Matrix deren i -ter Spaltenvektoren mit $\varphi(e_i) \in \mathbb{K}^m$ überein stimmt. Es gilt daher $\psi_A(e_i) = \varphi(e_i)$, für jedes $i = 1, \dots, n$. Ist nun $x \in \mathbb{K}^n$ mit Komponenten $x_i \in \mathbb{K}$, dann gilt offensichtlich $x = x_1 e_1 + \dots + x_n e_n$. Aus der Linearität von φ und ψ_A folgt daher

$$\begin{aligned} \varphi(x) &= \varphi(x_1 e_1 + \dots + x_n e_n) = x_1 \varphi(e_1) + \dots + x_n \varphi(e_n) \\ &= x_1 \psi_A(e_1) + \dots + x_n \psi_A(e_n) = \psi_A(x_1 e_1 + \dots + x_n e_n) = \psi_A(x). \end{aligned}$$

Da dies für jeden Vektor $x \in \mathbb{K}^n$ gilt, erhalten wir $\varphi = \psi_A$. Dies zeigt, dass die Zuordnung $M_{m \times n}(\mathbb{K}) \rightarrow L(\mathbb{K}^n, \mathbb{K}^m)$, $A \mapsto \psi_A$, auch surjektiv ist.

Aus Proposition II.4.3(c) erhalten wir $\psi_{A+A'}(x) = (A+A')x = Ax + A'x = \psi_A(x) + \psi_{A'}(x) = (\psi_A + \psi_{A'})(x)$, für jedes $x \in \mathbb{K}^n$, also $\psi_{A+A'} = \psi_A + \psi_{A'}$. Analog lässt sich $\psi_{\lambda A} = \lambda \psi_A$ nachrechnen. Somit ist $M_{m \times n}(\mathbb{K}) \rightarrow L(\mathbb{K}^n, \mathbb{K}^m)$, $A \mapsto \psi_A$, eine lineare Bijektion, also ein Isomorphismus, siehe Bemerkung II.3.11.

Aus Proposition II.4.3(a) erhalten wir auch $\psi_{AB}(x) = (AB)x = A(Bx) = \psi_A(Bx) = \psi_A(\psi_B(x)) = (\psi_A \circ \psi_B)(x)$, für jedes $x \in \mathbb{K}^k$, also $\psi_{AB} = \psi_A \circ \psi_B$. Analog folgt aus Proposition II.4.3(b) sofort $\psi_{I_n} = \text{id}_{\mathbb{K}^n}$. \square

II.4.5. BEMERKUNG. Nach Proposition II.4.3 bilden die quadratischen Matrizen, $M_{n \times n}(\mathbb{K})$, eine assoziative \mathbb{K} -Algebra mit Eins. Nach Satz II.4.4 ist der lineare Isomorphismus $M_{n \times n}(\mathbb{K}) \cong \text{end}(\mathbb{K}^n)$, $A \leftrightarrow \psi_A$, sogar ein Isomorphismus von \mathbb{K} -Algebren. Die Algebra $M_{n \times n}(\mathbb{K})$ ist i.A. nicht kommutativ, etwa gilt $\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \neq \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$, vgl. Bemerkung II.3.19. Im Fall $n = 1$ erhalten wir den Grundkörper, $M_{1 \times 1}(\mathbb{K}) = \mathbb{K}$.

Eine quadratische Matrix $A \in M_{n \times n}(\mathbb{K})$ wird *invertierbar* genannt, wenn eine Matrix $A' \in M_{n \times n}(\mathbb{K})$ existiert, sodass $AA' = I_n = A'A$. In diesem Fall ist die Matrix A' eindeutig bestimmt, denn aus $AA'' = I_n = A''A$ folgt mit Proposition II.4.3 $A'' = I_n A'' = (A'A)A'' = A'(AA'') = A'I_n = A'$. Diese eindeutig bestimmte Matrix wird *Inverse* von A genannt und von nun an mit A^{-1} bezeichnet, für eine invertierbare Matrix A gilt daher

$$AA^{-1} = I_n = A^{-1}A.$$

Mit A ist offensichtlich auch A^{-1} invertierbar,

$$(A^{-1})^{-1} = A.$$

Sind $A, B \in \text{GL}_n(\mathbb{K})$ zwei invertierbare Matrizen, dann ist auch AB invertierbar mit Inverser

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

denn $(AB)(B^{-1}A^{-1}) = AB B^{-1} A^{-1} = A I_n A^{-1} = A A^{-1} = I_n$ und analog erhalten wir $(B^{-1}A^{-1})(AB) = I_n$.

Die Menge aller invertierbaren $(n \times n)$ -Matrizen über \mathbb{K} wird mit $\mathrm{GL}_n(\mathbb{K})$ bezeichnet. Nach dem vorangehenden Absatz bildet $\mathrm{GL}_n(\mathbb{K})$ mit dem Matrizenprodukt eine Gruppe. Diese Gruppe ist i.A. nicht abelsch. Zum Beispiel sind die beiden Matrizen $A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ und $B = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$ invertierbar, $A^{-1} = \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix}$, $B^{-1} = \begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix}$, aber $AB = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix} \neq \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix} = BA$.

II.4.6. KOROLLAR. *Eine Matrix $A \in M_{n \times n}(\mathbb{K})$ ist genau dann invertierbar, wenn die lineare Abbildung, $\psi_A: \mathbb{K}^n \rightarrow \mathbb{K}^n$, $\psi_A(x) = Ax$, invertierbar ist und in diesem Fall gilt $\psi_A^{-1} = \psi_{A^{-1}}$. Der Isomorphismus aus Satz II.4.4 schränkt sich daher zu einem Gruppenisomorphismus $\mathrm{GL}_n(\mathbb{K}) \cong \mathrm{GL}(\mathbb{K}^n)$, $A \leftrightarrow \psi_A$, ein.*

BEWEIS. Ist $A \in M_{n \times n}(\mathbb{K})$ invertierbar, dann folgt $\psi_A \circ \psi_{A^{-1}} = \psi_{AA^{-1}} = \psi_{I_n} = \mathrm{id}_{\mathbb{K}^n}$ und analog $\psi_{A^{-1}} \circ \psi_A = \psi_{A^{-1}A} = \psi_{I_n} = \mathrm{id}_{\mathbb{K}^n}$. Somit ist $\psi_A: \mathbb{K}^n \rightarrow \mathbb{K}^n$ eine invertierbare lineare Abbildung mit Umkehrabbildung $\psi_A^{-1} = \psi_{A^{-1}}$. Sei nun umgekehrt $A \in M_{n \times n}(\mathbb{K})$, sodass $\psi_A: \mathbb{K}^n \rightarrow \mathbb{K}^n$ invertierbar ist. Nach Satz II.4.4 existiert daher $B \in M_{n \times n}(\mathbb{K})$ mit $\psi_A \circ \psi_B = \mathrm{id}_{\mathbb{K}^n} = \psi_B \circ \psi_A$. Wir erhalten $\psi_{AB} = \psi_A \circ \psi_B = \mathrm{id}_{\mathbb{K}^n} = \psi_{I_n}$ und analog $\psi_{BA} = \psi_B \circ \psi_A = \mathrm{id}_{\mathbb{K}^n} = \psi_{I_n}$. Aus Satz II.4.4 folgt somit $AB = I_n = BA$, also ist die Matrix A invertierbar. \square

II.4.7. BEISPIEL. Eine (1×1) -Matrix $A = (a) \in M_{1 \times 1}(\mathbb{K})$ ist genau dann invertierbar, wenn $a \neq 0$. In diesem Fall gilt $A^{-1} = (a^{-1})$.

II.4.8. BEISPIEL. Eine (2×2) -Matrix $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in M_{2 \times 2}(\mathbb{K})$ ist genau dann invertierbar, wenn $ad - bc \neq 0$. In diesem Fall gilt

$$A^{-1} = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}. \quad (\text{II.2})$$

Ist nämlich $ad - bc \neq 0$ dann gilt

$$\frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \frac{1}{ad - bc} \begin{pmatrix} da - bc & db - bd \\ -ca + ac & -cb + ad \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = I_2$$

und

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix} = \frac{1}{ad - bc} \begin{pmatrix} ad - bc & -ab + ba \\ cd - dc & -cb + da \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = I_2,$$

also ist A invertierbar mit Inverser (II.2). Sei nun umgekehrt $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ invertierbar mit Inverser $A^{-1} = \begin{pmatrix} a' & b' \\ c' & d' \end{pmatrix}$. Da

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = I_2 = AA^{-1} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} a' & b' \\ c' & d' \end{pmatrix} = \begin{pmatrix} aa' + bc' & ab' + bd' \\ ca' + dc' & cb' + dd' \end{pmatrix},$$

erhalten wir $aa' + bc' = 1$, $ab' + bd' = 0$, $ca' + dc' = 0$ und $cb' + dd' = 1$. Daraus folgt $(ad - bc)(a'd' - b'c') = (aa' + bc')(cb' + dd') = 1 \cdot 1 = 1$, also muss $ad - bc \neq 0$ gelten.

II.4.9. BEISPIEL. Wir wollen die Umkehrabbildung der linearen Abbildung $\varphi: \mathbb{K}^2 \rightarrow \mathbb{K}^2$, $\varphi \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} x_1 + 2x_2 \\ 3x_1 + 5x_2 \end{pmatrix}$, bestimmen, sofern diese existiert. Offensichtlich

gilt $\varphi \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = A \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ mit $A = \begin{pmatrix} 1 & 2 \\ 3 & 5 \end{pmatrix}$. Nach Beispiel II.4.8 ist A invertierbar mit Inverser $A^{-1} = \begin{pmatrix} -5 & 2 \\ 3 & -1 \end{pmatrix}$. Nach Korollar II.4.6 ist daher auch φ invertierbar mit Umkehrabbildung $\varphi^{-1} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = A^{-1} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} -5 & 2 \\ 3 & -1 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} -5y_1 + 2y_2 \\ 3y_1 - y_2 \end{pmatrix}$.

Wir wollen an dieser Stelle noch einfache Eigenschaften einer weiteren Operationen mit Matrizen zusammenstellen, deren Bedeutung aber erst später klar werden wird. Unter der *Transponierten* einer Matrix $A \in M_{m \times n}(\mathbb{K})$ verstehen wir jene Matrix $A^t \in M_{n \times m}(\mathbb{K})$, die wir durch Vertauschen von Zeilen und Spalten erhalten,

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix}, \quad A^t := \begin{pmatrix} a_{11} & \cdots & a_{m1} \\ \vdots & & \vdots \\ a_{1n} & \cdots & a_{nm} \end{pmatrix}.$$

In anderen Worten, $(A^t)_{ij} = A_{ji}$.

II.4.10. BEISPIEL. Etwa gilt

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

II.4.11. PROPOSITION (Eigenschaften der Transponierten). *Sei \mathbb{K} ein Körper und $n, m \in \mathbb{N}$. Dann ist die Abbildung $M_{m \times n}(\mathbb{K}) \rightarrow M_{n \times m}(\mathbb{K})$, $A \mapsto A^t$, linear, d.h. für alle $A, \tilde{A} \in M_{m \times n}(\mathbb{K})$ und $\lambda \in \mathbb{K}$ gilt*

$$(A + \tilde{A})^t = A^t + \tilde{A}^t \quad \text{sowie} \quad (\lambda A)^t = \lambda A^t.$$

Darüber hinaus haben wir stets

$$(AB)^t = B^t A^t, \quad (A^t)^t = A \quad \text{und} \quad I_n^t = I_n,$$

für beliebige $A \in M_{m \times n}(\mathbb{K})$ und $B \in M_{n \times l}(\mathbb{K})$. Eine quadratische Matrix $C \in M_{n \times n}(\mathbb{K})$ ist genau dann invertierbar, wenn ihre Transponierte invertierbar ist und in diesem Fall gilt

$$(C^t)^{-1} = (C^{-1})^t.$$

BEWEIS. Es gilt $(A + \tilde{A})^t = A^t + \tilde{A}^t$, denn $((A + \tilde{A})^t)_{ij} = (A + \tilde{A})_{ji} = A_{ji} + \tilde{A}_{ji} = (A^t)_{ij} + (\tilde{A}^t)_{ij} = (A^t + \tilde{A}^t)_{ij}$. Analog lässt sich $(\lambda A)^t = \lambda A^t$ zeigen, $((\lambda A)^t)_{ij} = (\lambda A)_{ji} = \lambda A_{ji} = \lambda (A^t)_{ij} = (\lambda A^t)_{ij}$. Schließlich gilt auch $(AB)^t = B^t A^t$, denn $((AB)^t)_{ij} = (AB)_{ji} = \sum_{k=1}^n A_{jk} B_{ki} = \sum_{k=1}^n B_{ki} A_{jk} = \sum_{k=1}^n (B^t)_{ik} (A^t)_{kj} = (B^t A^t)_{ij}$. Für eine invertierbare quadratische Matrix C folgt $C^t (C^{-1})^t = (C^{-1} C)^t = I_n^t = I_n$ und analog $(C^{-1})^t C^t = (C C^{-1})^t = I_n^t = I_n$, also ist auch C^t invertierbar mit Inverser $(C^t)^{-1} = (C^{-1})^t$. Die restlichen Behauptungen sind trivial. \square

II.4.12. BEISPIEL. Die *symmetrischen Matrizen*, $\{A \in M_{n \times n}(\mathbb{K}) : A^t = A\}$, bilden einen Teilraum von $M_{n \times n}(\mathbb{K})$, denn diese Teilmenge stimmt offensichtlich mit dem Kern der linearen Abbildung $M_{n \times n}(\mathbb{K}) \rightarrow M_{n \times n}(\mathbb{K})$, $A \mapsto A^t - A$,

überein. Auch die *schiefssymmetrischen Matrizen*, $\{A \in M_{n \times n}(\mathbb{K}) : A^t = -A\}$, bilden einen Teilraum von $M_{n \times n}(\mathbb{K})$, denn dies ist genau der Kern der linearen Abbildung $M_{n \times n}(\mathbb{K}) \rightarrow M_{n \times n}(\mathbb{K})$, $A \mapsto A^t + A$.

II.4.13. BEMERKUNG (Matrizen und Gleichungssysteme). Wir wollen uns nun überlegen wie sich Fragen zur Lösbarkeit eines linearen Gleichungssystems mit Matrizen bzw. linearen Abbildungen formulieren lassen. Unter der *Koeffizientenmatrix* eines linearen Gleichungssystems

$$\begin{aligned} a_{11}x_1 + \cdots + a_{1n}x_n &= y_1 \\ &\vdots \\ a_{m1}x_1 + \cdots + a_{mn}x_n &= y_m \end{aligned}$$

verstehen wir die Matrix $A \in M_{m \times n}(\mathbb{K})$ mit Eintragungen $A_{ij} := a_{ij}$. Das Gleichungssystem lässt sich damit in kompakter Form schreiben,

$$Ax = y$$

wobei $x \in \mathbb{K}^n$ und $y \in \mathbb{K}^m$. Die mit A assoziierte lineare Abbildung

$$\psi_A: \mathbb{K}^n \rightarrow \mathbb{K}^m, \quad \psi_A(x) = Ax,$$

liefert die linke Seite des Gleichungssystems,

$$\psi_A \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} a_{11}x_1 + \cdots + a_{1n}x_n \\ \vdots \\ a_{m1}x_1 + \cdots + a_{mn}x_n \end{pmatrix} = x_1 \begin{pmatrix} a_{11} \\ \vdots \\ a_{m1} \end{pmatrix} + \cdots + x_n \begin{pmatrix} a_{1n} \\ \vdots \\ a_{mn} \end{pmatrix}.$$

Das Bild der linearen Abbildung ψ_A stimmt daher mit dem Teilraum aller $y \in \mathbb{K}^m$ überein, für die das Gleichungssystem $Ax = y$ lösbar ist,

$$\text{img}(\psi_A) = \{y \in \mathbb{K}^m \mid \exists x \in \mathbb{K}^n : Ax = y\}.$$

Die lineare Abbildung ψ_A ist also genau dann surjektiv, wenn das Gleichungssystem $Ax = y$ für jedes $y \in \mathbb{K}^m$ mindestens eine Lösung $x \in \mathbb{K}^n$ hat.

Die Abbildung ψ_A ist genau dann injektiv, wenn das Gleichungssystem $Ax = y$ für jedes $y \in \mathbb{K}^m$ höchstens eine Lösung $x \in \mathbb{K}^n$ besitzt. Nach Proposition II.3.22 ist dies genau dann der Fall, wenn ψ_A trivialen Kern hat, d.h. wenn $\ker(\psi_A) = \{0\}$ gilt. Dies wiederum bedeutet gerade, dass das Gleichungssystem $Ax = 0$ nur die triviale Lösung $x = 0$ besitzt. Die lineare Abbildung ψ_A ist genau dann bijektiv, wenn das Gleichungssystem $Ax = y$ für jedes $y \in \mathbb{K}^m$ genau eine Lösung $x \in \mathbb{K}^n$ hat. Wir werden später sehen, dass dies nur im Fall $n = m$ möglich ist. Die linearen Isomorphismen $\mathbb{K}^n \rightarrow \mathbb{K}^m$ entsprechen daher genau den eindeutig lösbar Gleichungssystemen bzw. den invertierbaren Matrizen.

Das Gleichungssystem

$$Ax = 0$$

wird als das mit $Ax = y$ assoziierte *homogene Gleichungssystem* bezeichnet. Seine Lösungsmenge stimmt mit dem Kern von ψ_A , d.h. dem Teilraum

$$\ker(\psi_A) = \{x \in \mathbb{K}^n : Ax = 0\}$$

überein. Ist $\xi \in \mathbb{K}^n$ eine Lösung des linearen Gleichungssystems, d.h. $A\xi = y$, dann gilt

$$\{x \in \mathbb{K}^n : Ax = y\} = \{\xi + x : Ax = 0\} = \xi + \ker(\psi_A),$$

wir erhalten daher alle Lösungen von $Ax = y$ indem wir zu einer speziellen Lösung ξ , d.h. $A\xi = y$, alle Lösungen des homogenen Systems $Ax = 0$ addieren.

II.5. Summen und Komplemente. Sind W_1 und W_2 zwei Teilräume eines Vektorraums V , dann bildet deren Vereinigung, $W_1 \cup W_2$, i.A. keinen Teilraum von V , vgl. Übungsaufgabe 19. Wir werden nun den kleinsten Teilraum von V betrachten, der $W_1 \cup W_2$ enthält.

II.5.1. DEFINITION (Summe von Teilräumen). Sind W_1 und W_2 zwei Teilräume eines Vektorraums V , so wird

$$W_1 + W_2 := \{w_1 + w_2 \mid w_1 \in W_1, w_2 \in W_2\}$$

die *Summe* der Teilräume W_1 und W_2 genannt.

II.5.2. PROPOSITION. Sind W_1 und W_2 zwei Teilräume eines Vektorraums V , dann ist $W_1 + W_2$ der kleinste Teilraum von V , der $W_1 \cup W_2$ enthält. D.h. $W_1 + W_2$ ist ein Teilraum von V , es gilt $W_1 \cup W_2 \subseteq W_1 + W_2$ und für jeden weiteren Teilraum W von V mit $W_1 \cup W_2 \subseteq W$ gilt schon $W_1 + W_2 \subseteq W$.

BEWEIS. Da $0 \in W_2$ gilt $W_1 \subseteq W_1 + W_2$, denn jedes $w_1 \in W_1$ lässt sich in der Form $w_1 = w_1 + 0$ schreiben. Analog haben wir auch $W_2 \subseteq W_1 + W_2$. Zusammen erhalten wir $W_1 \cup W_2 \subseteq W_1 + W_2$. Insbesondere ist $W_1 + W_2$ nicht leer. Um die Abgeschlossenheit unter Addition zu zeigen betrachten wir zwei Elemente $w, w' \in W_1 + W_2$. Nach Definition der Summe existieren daher $w_1, w'_1 \in W_1$ und $w_2, w'_2 \in W_2$ mit $w = w_1 + w_2$ und $w' = w'_1 + w'_2$. Da W_1 und W_2 Teilräume sind, folgt

$$w + w' = (w_1 + w_2) + (w'_1 + w'_2) = \underbrace{(w_1 + w'_1)}_{\in W_1} + \underbrace{(w_2 + w'_2)}_{\in W_2},$$

also liegt auch $w + w'$ in $W_1 + W_2$. Somit ist $W_1 + W_2$ abgeschlossen unter Addition. Seien nun $\lambda \in \mathbb{K}$ und $w \in W_1 + W_2$. Es existieren daher $w_1 \in W_1$ und $w_2 \in W_2$ mit $w = w_1 + w_2$. Aufgrund von $\lambda w = \lambda(w_1 + w_2) = \lambda w_1 + \lambda w_2$ liegt also auch λw in $W_1 + W_2$. Somit ist $W_1 + W_2$ auch abgeschlossen unter Skalarmultiplikation und daher ein Teilraum von V . Die Minimalität von $W_1 + W_2$ ist offensichtlich, jeder Teilraum der W_1 und W_2 enthält muss auch alle Vektoren der Form $w_1 + w_2$ mit $w_1 \in W_1, w_2 \in W_2$ enthalten. \square

II.5.3. BEMERKUNG. Sind W , W_1 , W_2 und W_3 Teilräume eines Vektorraums V , dann gelten offensichtlich die Relationen $W_1 + W_2 = W_2 + W_1$, $W_1 + (W_2 + W_3) = (W_1 + W_2) + W_3$, und $W + \{0\} = W$. Weiters gilt $W_1 + W_2 = W_2$ genau dann, wenn $W_1 \subseteq W_2$, vgl. Übungsaufgabe 46.

Nach Definition der Summe lässt sich jedes Element $v \in W_1 + W_2$ in der Form $v = w_1 + w_2$ schreiben, $w_1 \in W_1$, $w_2 \in W_2$. Wie das folgende Beispiel zeigt ist diese Darstellung i.A. jedoch nicht eindeutig.

II.5.4. BEISPIEL. Betrachte folgende beiden Teilräume von \mathbb{K}^3 ,

$$W_1 = \left\{ \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \in \mathbb{K}^3 : x_1 = 0 \right\} \quad \text{und} \quad W_2 = \left\{ \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \in \mathbb{K}^3 : x_2 = 0 \right\}.$$

Offensichtlich gilt $\mathbb{K}^3 = W_1 + W_2$, denn jeder Vektor lässt sich in der Form $\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ x_2 \\ x_3 \end{pmatrix} + \begin{pmatrix} x_1 \\ 0 \\ 0 \end{pmatrix}$ schreiben, wobei der erste Summand in W_1 und der zweite in W_2 liegt. Allerdings ist diese Darstellung nicht eindeutig, $\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ x_2 \\ x_3 - 1 \end{pmatrix} + \begin{pmatrix} x_1 \\ 0 \\ 1 \end{pmatrix}$ ist eine andere Zerlegung mit den selben Eigenschaften.

II.5.5. PROPOSITION. Sind W_1 und W_2 zwei Teilräume eines Vektorraums V , dann sind die folgenden Aussagen äquivalent:

- (a) $W_1 + W_2 = V$ und $W_1 \cap W_2 = \{0\}$.
- (b) Jedes $v \in V$ lässt sich auf eindeutige Weise in der Form $v = w_1 + w_2$ schreiben, für gewisse $w_1 \in W_1$ und $w_2 \in W_2$.
- (c) Zu je zwei linearen Abbildungen $\varphi_1: W_1 \rightarrow U$ und $\varphi_2: W_2 \rightarrow U$ existiert genau eine lineare Abbildung $\varphi: V \rightarrow U$, sodass $\varphi|_{W_1} = \varphi_1$ und $\varphi|_{W_2} = \varphi_2$.

BEWEIS. Ad (a) \Rightarrow (b): Da $V = W_1 + W_2$ lässt sich jedes $v \in V$ in der Form $v = w_1 + w_2$ für gewisse $w_1 \in W_1$ und $w_2 \in W_2$ schreiben. Sei nun $v = w'_1 + w'_2$ eine weiter solche Darstellung, d.h. $w'_1 \in W_1$ und $w'_2 \in W_2$. Es folgt $w_1 + w_2 = v = w'_1 + w'_2$ und daher

$$w_1 - w'_1 = w'_2 - w_2.$$

Beachte, dass die linke Seite dieser Gleichung in W_1 und die rechte Seite in W_2 liegt. Beide Seiten liegen daher in $W_1 \cap W_2 = \{0\}$. Wir erhalten also $w_1 - w'_1 = 0$ und $w'_2 - w_2 = 0$, d.h. $w_1 = w'_1$ und $w_2 = w'_2$. Somit sind w_1 und w_2 in der Darstellung $v = w_1 + w_2$ eindeutig bestimmt.

Ad (b) \Rightarrow (c): Seien also $\varphi_1: W_1 \rightarrow U$ und $\varphi_2: W_2 \rightarrow U$ zwei lineare Abbildungen. Wir definieren eine Abbildung $\varphi: V \rightarrow U$ durch $\varphi(v) := \varphi_1(w_1) + \varphi_2(w_2)$, wobei $w_1 \in W_1$ und $w_2 \in W_2$ jene eindeutig bestimmten Vektoren bezeichnen, für die $v = w_1 + w_2$ gilt. Offensichtlich gilt $\varphi|_{W_1} = \varphi_1$ und $\varphi|_{W_2} = \varphi_2$. Wir zeigen nun, dass φ linear ist. Sei dazu $v' \in V$ und $v' = w'_1 + w'_2$ die eindeutige Zerlegung mit $w'_1 \in W_1$ und $w'_2 \in W_2$. Für die Zerlegung der Summe, $v + v'$, ergibt sich

$v + v' = (w_1 + w'_1) + (w_2 + w'_2)$, wobei $w_1 + w'_1 \in W_1$ und $w_2 + w'_2 \in W_2$. Es folgt

$$\begin{aligned}\varphi(v + v') &= \varphi_1(w_1 + w'_1) + \varphi_2(w_2 + w'_2) \\ &= (\varphi_1(w_1) + \varphi_1(w'_1)) + (\varphi_2(w_2) + \varphi_2(w'_2)) \\ &= (\varphi_1(w_1) + \varphi_2(w_2)) + (\varphi_1(w'_1) + \varphi_2(w'_2)) \\ &= \varphi(v) + \varphi(v').\end{aligned}$$

Ist $\lambda \in \mathbb{K}$, so gilt $\lambda v = \lambda w_1 + \lambda w_2$ mit $\lambda w_1 \in W_1$ und $\lambda w_2 \in W_2$ und daher $\varphi(\lambda v) = \varphi_1(\lambda w_1) + \varphi_2(\lambda w_2) = \lambda \varphi_1(w_1) + \lambda \varphi_2(w_2) = \lambda(\varphi_1(w_1) + \varphi_2(w_2)) = \lambda \varphi(v)$. Dies zeigt, dass φ tatsächlich eine lineare Abbildung darstellt. Die Eindeutigkeit von φ ist offensichtlich, jedes $v \in V$ lässt sich ja in der Form $v = w_1 + w_2$, $w_1 \in W_1$, $w_2 \in W_2$, schreiben, für ein lineares $\varphi: V \rightarrow U$ mit $\varphi|_{W_1} = \varphi_1$ und $\varphi|_{W_2} = \varphi_2$ folgt daher $\varphi(v) = \varphi(w_1 + w_2) = \varphi(w_1) + \varphi(w_2) = \varphi|_{W_1}(w_1) + \varphi|_{W_2}(w_2) = \varphi_1(w_1) + \varphi_2(w_2)$.

Ad (c) \Rightarrow (a): Nach Voraussetzung existiert eine lineare Abbildung $\pi_1: V \rightarrow W_1$ mit $\pi_1|_{W_1} = \text{id}_{W_1}$ und $\pi_1|_{W_2} = 0$. Daraus erhalten wir sofort $W_1 \cap W_2 = \{0\}$, denn für $v \in W_1 \cap W_2$ folgt $v = \pi_1(v) = 0$. Es existiert aber auch eine lineare Abbildung $\pi_2: V \rightarrow W_2$ mit $\pi_2|_{W_1} = 0$ und $\pi_2|_{W_2} = \text{id}_{W_2}$. Fassen wir π_1 und π_2 als lineare Abbildungen $\pi_1: V \rightarrow V$ und $\pi_2: V \rightarrow V$ auf, dann gilt für ihre Summe, $\pi_1 + \pi_2: V \rightarrow V$ nun $(\pi_1 + \pi_2)|_{W_1} = \text{id}_V|_{W_1}$ und $(\pi_1 + \pi_2)|_{W_2} = \text{id}_V|_{W_2}$. Aus der Eindeutigkeitsaussage in (c) folgt daher $\pi_1 + \pi_2 = \text{id}_V$. Für jedes $v \in V$ erhalten wir somit $v = \text{id}_V(v) = \pi_1(v) + \pi_2(v)$ mit $\pi_1(v) \in W_1$ und $\pi_2(v) \in W_2$, also $V = W_1 + W_2$. \square

II.5.6. DEFINITION (Innere direkte Summe und Komplement). Zwei Teilräume W_1 und W_2 eines Vektorraums V heißen *komplementär*, falls $W_1 + W_2 = V$ und $W_1 \cap W_2 = \{0\}$ gilt. In diesem Fall sagen wir V ist die (*innere*) *direkte Summe* von W_1 und W_2 und notieren dies durch $V = W_1 \oplus W_2$. Auch wird W_2 als *ein Komplement* von W_1 in V bezeichnet. Die nach Proposition II.5.5 eindeutig bestimmte lineare Abbildung $\pi_1: V \rightarrow W_1$ mit $\pi_1|_{W_1} = \text{id}_{W_1}$ und $\pi_1|_{W_2} = 0$ wird als *Projektion auf W_1 längs W_2* bezeichnet. Mit vertauschten Rollen wird auch W_1 ein Komplement von W_2 in V genannt. Die eindeutig bestimmte lineare Abbildung $\pi_2: V \rightarrow W_2$ mit $\pi_2|_{W_2} = \text{id}_{W_2}$ und $\pi_2|_{W_1} = 0$ wird als *Projektion auf W_2 längs W_1* bezeichnet. Auch wird $\pi_2 = \text{id}_V - \pi_1$ die zu π_1 *komplementäre Projektion* genannt.

Seien W_1 und W_2 zwei komplementäre Teilräume eines Vektorraums V , d.h. $W_1 \oplus W_2 = V$, und $\pi_1: V \rightarrow W_1$ sowie $\pi_2: V \rightarrow W_2$ die damit assoziierten Projektionen. Ist $v \in V$ und $v = w_1 + w_2$ die eindeutige Zerlegung mit $w_1 \in W_1$ und $w_2 \in W_2$, dann gilt $\pi_1(v) = \pi_1(w_1 + w_2) = \pi_1(w_1) + \pi_1(w_2) = w_1 + 0 = w_1$ und analog $\pi_2(v) = w_2$. Die beiden Projektionen liefern uns also für jedes $v \in V$ die eindeutige Zerlegung $v = w_1 + w_2$ mit $w_1 = \pi_1(v) \in W_1$ und $w_2 = \pi_2(v) \in W_2$. Fassen wir diese Projektionen als lineare Abbildungen $\pi_1, \pi_2: V \rightarrow V$ auf, dann folgt $\pi_1 \circ \pi_1 = \pi_1$, $\pi_2 \circ \pi_2 = \pi_2$, $\pi_1 + \pi_2 = \text{id}_V$ und $\pi_1 \circ \pi_2 = 0 = \pi_2 \circ \pi_1$.

II.5.7. DEFINITION (Projektor). Unter einem Projektor verstehen wir eine lineare Abbildung $\pi: V \rightarrow V$ für die $\pi \circ \pi = \pi$ gilt.

II.5.8. PROPOSITION. Ist $\pi: V \rightarrow V$ ein Projektor, dann gilt

$$V = \text{img}(\pi) \oplus \ker(\pi)$$

und π stimmt mit der Projektion auf $\text{img}(\pi)$ längs $\ker(\pi)$ überein. Auch $\pi' := \text{id}_V - \pi$ ist ein Projektor, er stimmt mit der komplementären Projektion auf $\ker(\pi)$ längs $\text{img}(\pi)$ überein. Weiters haben wir $\text{img}(\pi) = \{v \in V : \pi(v) = v\} = \ker(\pi')$, $\text{img}(\pi') = \{v \in V : \pi'(v) = v\} = \ker(\pi)$ und es gelten die Formeln

$$\pi \circ \pi = \pi, \quad \pi' \circ \pi' = \pi', \quad \pi + \pi' = \text{id}_V \quad \text{sowie} \quad \pi \circ \pi' = 0 = \pi' \circ \pi.$$

BEWEIS. Zunächst ist auch $\pi' = \text{id}_V - \pi: V \rightarrow V$ ein Projektor, denn

$$\begin{aligned} \pi' \circ \pi' &= (\text{id}_V - \pi) \circ (\text{id}_V - \pi) \\ &= \text{id}_V \circ \text{id}_V - \pi \circ \text{id}_V - \text{id}_V \circ \pi + \pi \circ \pi \\ &= \text{id}_V - \pi - \pi + \pi \\ &= \text{id}_V - \pi = \pi'. \end{aligned}$$

Auch erhalten wir sofort $\pi \circ \pi' = \pi \circ (\text{id}_V - \pi) = \pi \circ \text{id}_V - \pi \circ \pi = \pi - \pi = 0$ und analog $\pi' \circ \pi = 0$. Die Relationen $\text{img}(\pi) \supseteq \{v \in V : \pi(v) = v\} = \ker(\pi')$ sind offensichtlich. Ist $v \in \text{img}(\pi)$, dann existiert $w \in V$ mit $v = \pi(w)$ und wir erhalten $\pi(v) = \pi(\pi(w)) = (\pi \circ \pi)(w) = \pi(w) = v$. Dies zeigt $\text{img}(\pi) \subseteq \{v \in V : \pi(v) = v\}$, es gilt daher $\text{img}(\pi) = \{v \in V : \pi(v) = v\} = \ker(\pi')$. Wenden wir dies auf den Projektor π' an, erhalten wir auch $\text{img}(\pi') = \{v \in V : \pi'(v) = v\} = \ker(\pi)$. Daraus folgt nun $\text{img}(\pi) \cap \ker(\pi) = \{0\}$, denn für $v \in \text{img}(\pi) \cap \ker(\pi)$ gilt $v = \pi(v) = 0$. Schließlich lässt sich jedes $v \in V$ in der Form $v = \pi(v) + \pi'(v)$ schreiben, wobei $\pi(v) \in \text{img}(\pi)$ und $\pi'(v) \in \text{img}(\pi') = \ker(\pi)$. Dies zeigt $V = \text{img}(\pi) + \ker(\pi)$, also $V = \text{img}(\pi) \oplus \ker(\pi)$. \square

II.5.9. BEMERKUNG (Spiegelungen). Sei $V = W_+ \oplus W_-$. Nach Proposition II.5.5 existiert eine eindeutige lineare Abbildung $\sigma: V \rightarrow V$, sodass $\sigma(v) = v$ für alle $v \in W_+$, und $\sigma(v) = -v$ für alle $v \in W_-$. Diese Abbildung σ wird *Spiegelung an W_+ längs W_-* genannt. Beachte $\sigma \circ \sigma = \text{id}_V$. Die komplementäre Spiegelung an W_- längs W_+ ist durch $-\sigma$ gegeben. Bezeichnen $\pi_+: V \rightarrow V$ und $\pi_-: V \rightarrow V$ die mit der Zerlegung $V = W_- \oplus W_+$ assoziierten Projektoren, dann gilt $\sigma = \pi_+ - \pi_-$. Ist $2 \neq 0 \in \mathbb{K}$, dann lassen sich auch die Projektionen durch die Spiegelungen ausdrücken, $\pi_+ = \frac{1}{2}(\text{id}_V + \sigma)$ sowie $\pi_- = \frac{1}{2}(\text{id}_V - \sigma)$, siehe auch Übungsaufgabe 50.

II.5.10. BEISPIEL. Betrachte die beiden Teilräume von \mathbb{R}^3 ,

$$E := \left\{ \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} : 2x_1 + 3x_2 + 4x_3 = 0 \right\} \quad \text{und} \quad G := \left\{ \lambda \begin{pmatrix} 5 \\ 6 \\ 7 \end{pmatrix} : \lambda \in \mathbb{R} \right\}.$$

Wir wollen uns nun davon überzeugen, dass \mathbb{R}^3 innere direkte Summe von E und G ist, d.h. $\mathbb{R}^3 = E \oplus G$. Zunächst gilt $E \cap G = \{0\}$, denn liegt ein Element

$x = \lambda \begin{pmatrix} 5 \\ 6 \\ 7 \end{pmatrix} = \begin{pmatrix} 5\lambda \\ 6\lambda \\ 7\lambda \end{pmatrix}$ aus G auch in E , so folgt $0 = 2 \cdot 5\lambda + 3 \cdot 6\lambda + 4 \cdot 7\lambda = 56\lambda$, also $\lambda = 0$ und damit $x = 0$. Andererseits haben wir auch $E + G = \mathbb{R}^3$, denn jeder Vektor aus \mathbb{R}^3 lässt sich in der Form

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \underbrace{\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}}_{\in E} - \lambda \underbrace{\begin{pmatrix} 5 \\ 6 \\ 7 \end{pmatrix}}_{\in G} + \lambda \underbrace{\begin{pmatrix} 5 \\ 6 \\ 7 \end{pmatrix}}_{\in G}, \quad \lambda = \frac{2x_1 + 3x_2 + 4x_3}{56},$$

schreiben, wobei λ so gewählt wurde, dass der erste Summand tatsächlich in E liegt. Dies zeigt $\mathbb{R}^3 = E \oplus G$. Für die Projektion auf G längs E , $\pi_G: \mathbb{R}^3 \rightarrow \mathbb{R}^3$, erhalten wir

$$\pi_G \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \frac{2x_1 + 3x_2 + 4x_3}{56} \begin{pmatrix} 5 \\ 6 \\ 7 \end{pmatrix} = \frac{1}{56} \begin{pmatrix} 10 & 15 & 20 \\ 12 & 18 & 24 \\ 14 & 21 & 28 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$$

Für die Projektion auf E längs G , $\pi_E: \mathbb{R}^3 \rightarrow \mathbb{R}^3$, folgt

$$\pi_E \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} - \frac{2x_1 + 3x_2 + 4x_3}{56} \begin{pmatrix} 5 \\ 6 \\ 7 \end{pmatrix} = \frac{1}{56} \begin{pmatrix} 46 & -15 & -20 \\ -12 & 38 & -24 \\ -14 & -21 & 28 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$$

Die Spiegelung an E längs G , d.h. $\sigma = \pi_E - \pi_G: \mathbb{R}^3 \rightarrow \mathbb{R}^3$, ist daher durch

$$\sigma \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \frac{1}{56} \begin{pmatrix} 36 & -30 & -40 \\ -24 & 20 & -48 \\ -28 & -42 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix},$$

gegeben, siehe Bemerkung II.5.9 und Übungsaufgabe 48.

II.5.11. BEISPIEL. Es bezeichne $X \subseteq \mathbb{R}$ ein um Null symmetrisches Intervall, etwa $X = (-a, a)$ mit $0 < a \leq \infty$. Wir wollen nun zeigen, dass $V := F(X, \mathbb{R})$, der Vektorraum aller Funktionen $X \rightarrow \mathbb{R}$, innere direkte Summe des Teilraums aller *geraden Funktionen*,

$$G := \{f \in F(X, \mathbb{R}) \mid \forall x \in X : f(-x) = f(x)\},$$

und des Teilraums aller *ungeraden Funktionen*,

$$U := \{f \in F(X, \mathbb{R}) \mid \forall x \in X : f(-x) = -f(x)\},$$

ist, d.h. $V = G \oplus U$. In Beispiel II.2.14 haben wir bereits verifiziert, dass G und U Teilräume von $F(X, \mathbb{R})$ bilden, für die $G \cap U = \{0\}$ gilt. Andererseits lässt sich jede Funktion $f: X \rightarrow \mathbb{R}$ in der Form $f = g + u$ schreiben, wobei $g: X \rightarrow \mathbb{R}$, $g(x) := \frac{1}{2}(f(x) + f(-x))$, $u: X \rightarrow \mathbb{R}$, $u(x) := \frac{1}{2}(f(x) - f(-x))$, und für diese Summanden gilt $g \in G$ sowie $u \in U$. Dies zeigt $V = G \oplus U$. Für die Projektion auf G längs U , $\pi_G: V \rightarrow V$, und die Projektion auf U längs G , $\pi_U: V \rightarrow V$, erhalten wir

$$(\pi_G(f))(x) = \frac{1}{2}(f(x) + f(-x)) \quad \text{und} \quad (\pi_U(f))(x) = \frac{1}{2}(f(x) - f(-x)).$$

Die Projektion π_G macht Funktionen gerade und zwar so, dass bereits gerade Funktionen unverändert bleiben und ungerade Funktionen auf 0 abgebildet werden. Analog macht die Projektion π_U Funktionen ungerade, wobei bereits ungerade Funktionen unverändert bleiben und gerade Funktionen auf 0 abgebildet werden. Beachte auch, dass die Spiegelung an G längs U , d.h. $\sigma = \pi_G - \pi_U: F(X, \mathbb{R}) \rightarrow F(X, \mathbb{R})$, durch $\sigma(f)(x) = f(-x)$ gegeben ist, $x \in X$. Mit Hilfe der Involution $\nu: X \rightarrow X$, $\nu(x) := -x$, lässt sich dies auch als $\sigma(f) = f \circ \nu$ schreiben.

II.5.12. BEISPIEL. Sei \mathbb{K} ein Körper in dem $2 \neq 0$ gilt. In Beispiel II.4.12 haben wir gesehen, dass die symmetrischen bzw. schiefsymmetrischen Matrizen,

$$W := \{A \in M_{n \times n}(\mathbb{K}) : A^t = A\} \quad \text{und} \quad W' := \{A \in M_{n \times n}(\mathbb{K}) : A^t = -A\}$$

jeweils Teilräume von $M_{n \times n}(\mathbb{K})$ bilden. Wir wollen nun zeigen, dass $M_{n \times n}(\mathbb{K})$ direkte Summe dieser Teilräume ist, d.h.

$$M_{n \times n}(\mathbb{K}) = W \oplus W'.$$

Zunächst gilt $W \cap W' = \{0\}$, denn für $A \in W \cap W'$ folgt $A = A^t = -A$, also $2A = 0$ und daher $A = 0$. Andererseits lässt sich jede Matrix $A \in M_{n \times n}(\mathbb{K})$ in der Form

$$A = \frac{1}{2}(A + A^t) + \frac{1}{2}(A - A^t)$$

schreiben, wobei der erste Summand offensichtlich in W liegt, denn $(\frac{1}{2}(A + A^t))^t = \frac{1}{2}(A^t + A^{tt}) = \frac{1}{2}(A^t + A) = \frac{1}{2}(A + A^t)$, und der zweite Summand in W' liegt, denn $(\frac{1}{2}(A - A^t))^t = \frac{1}{2}(A^t - A^{tt}) = \frac{1}{2}(A^t - A) = -\frac{1}{2}(A - A^t)$. Dies zeigt $M_{n \times n}(\mathbb{K}) = W + W'$, also $M_{n \times n}(\mathbb{K}) = W \oplus W'$. Für die Projektion $\pi: M_{n \times n}(\mathbb{K}) \rightarrow W \subseteq M_{n \times n}(\mathbb{K})$ auf W längs W' bzw. die Projektion $\pi': M_{n \times n}(\mathbb{K}) \rightarrow W' \subseteq M_{n \times n}(\mathbb{K})$ auf W' längs W erhalten wir

$$\pi(A) = \frac{1}{2}(A + A^t) \quad \text{und} \quad \pi'(A) = \frac{1}{2}(A - A^t).$$

Die Spiegelung an W längs W' , d.h. $\sigma = \pi - \pi': M_{n \times n}(\mathbb{K}) \rightarrow M_{n \times n}(\mathbb{K})$, ist daher durch $\sigma(A) = A^t$ gegeben.

II.5.13. PROPOSITION. Sei $\varphi: V \rightarrow W$ eine surjektive lineare Abbildung und V' ein zu $\ker(\varphi)$ komplementärer Teilraum in V , d.h. $V = \ker(\varphi) \oplus V'$. Dann ist die Einschränkung $\varphi|_{V'}: V' \rightarrow W$ ein linearer Isomorphismus.

BEWEIS. Zunächst ist $\varphi|_{V'}$ injektiv, da $\ker(\varphi|_{V'}) = \{v' \in V' : \varphi(v') = 0\} = \ker(\varphi) \cap V' = \{0\}$, siehe auch Proposition II.3.22. Um auch die Surjektivität von $\varphi|_{V'}$ zu zeigen, sei $w \in W$ beliebig. Wegen der Surjektivität von φ gibt es $v \in V$ mit $\varphi(v) = w$. Da $V = \ker(\varphi) + V'$, existiert $v' \in V'$, sodass $v - v' \in \ker(\varphi)$. Es folgt $w = \varphi(v) = \varphi(v - v' + v') = \varphi(v - v') + \varphi(v') = 0 + \varphi(v') = \varphi(v')$, also $w \in \varphi(V')$. Somit ist $\varphi|_{V'}$ eine lineare Bijektion, also ein Isomorphismus. \square

II.5.14. BEMERKUNG. Betrachte ein lineares Gleichungssystem $Ax = y$, wobei $A \in M_{m \times n}(\mathbb{K})$, $x \in \mathbb{K}^n$ und $y \in \mathbb{K}^m$. Weiters bezeichne $\psi: \mathbb{K}^n \rightarrow \mathbb{K}^m$,

$\psi(x) = Ax$, die damit assoziierte lineare Abbildung, und $W := \text{img}(\psi)$ den Teilraum aller $y \in \mathbb{K}^m$, für die das Gleichungssystem $Ax = y$ lösbar ist. Gelingt es einen zu $\ker(\psi)$ komplementären Teilraum V' zu bestimmen, d.h. $\ker(\psi) \oplus V' = V$, dann ist die Einschränkung $\psi|_{V'}: V' \rightarrow W$ nach Proposition II.5.13 ein Isomorphismus, und ihre Umkehrabbildung, $\xi := \psi|_{V'}^{-1}: W \rightarrow V' \subseteq V$, liefert dann zu jedem $y \in W$ eine spezielle Lösung $\xi(y) \in \mathbb{K}^n$, die linear von y abhängt.

II.6. Quotientenräume. Es sei V ein Vektorraum über \mathbb{K} und $W \subseteq V$ ein Teilraum. Wir definieren auf V nun eine Relation \sim durch

$$v \sim v' :\Leftrightarrow v' - v \in W.$$

II.6.1. LEMMA. *Diese Relation \sim ist eine Äquivalenzrelation auf V .*

BEWEIS. Die Relation ist reflexiv, denn für jedes $v \in V$ gilt $v - v = 0 \in W$, also $v \sim v$. Die Relation ist symmetrisch, denn aus $v \sim v'$ folgt $v' - v \in W$, somit ist auch $v - v' = -(v' - v) \in W$ und daher $v' \sim v$. Die Relation ist transitiv, denn aus $v \sim v'$ und $v' \sim v''$ erhalten wir $v' - v \in W$ und $v'' - v' \in W$, somit auch $v'' - v = (v'' - v') + (v' - v) \in W$ und daher $v \sim v''$. \square

II.6.2. LEMMA. *Sind $v, v_1, v_2, v', v'_1, v'_2 \in V$ und $\lambda \in \mathbb{K}$, dann gilt:*

- (a) *Aus $v_1 \sim v'_1$ und $v_2 \sim v'_2$ folgt $(v_1 + v_2) \sim (v'_1 + v'_2)$.*
 (b) *Aus $v \sim v'$ folgt $\lambda v \sim \lambda v'$.*

BEWEIS. Ad (a): Gilt $v_1 \sim v'_1$ und $v_2 \sim v'_2$ so folgt $v'_1 - v_1 \in W$ und $v'_2 - v_2 \in W$, somit auch $(v'_1 + v'_2) - (v_1 + v_2) = (v'_1 - v_1) + (v'_2 - v_2) \in W$ und daher $(v_1 + v_2) \sim (v'_1 + v'_2)$. Ad (b): Aus $v \sim v'$ erhalten wir $v' - v \in W$, somit auch $\lambda v' - \lambda v = \lambda(v' - v) \in W$ und daher $\lambda v \sim \lambda v'$. \square

Die Äquivalenzklassen von \sim , sind genau die Translate von W . Wir werden die von $v \in V$ repräsentierte Äquivalenzklasse meist mit

$$[v] := \{v' \in V : v \sim v'\} = \{v + w : w \in W\} = v + W$$

bezeichnen. Für die Menge der Äquivalenzklassen schreiben wir V/W . Wir definieren nun auf V/W Addition $V/W \times V/W \xrightarrow{+} V/W$ und Skalarmultiplikation $\mathbb{K} \times V/W \xrightarrow{\cdot} V/W$ durch

$$[v_1] + [v_2] := [v_1 + v_2] \quad \text{und} \quad \lambda[v] := [\lambda v],$$

wobei $v, v_1, v_2 \in V$ und $\lambda \in \mathbb{K}$. *Beachte, dass diese Operationen nach Lemma II.6.2 tatsächlich wohldefiniert sind!* Wir werden uns nun davon überzeugen, dass V/W dadurch zu einem \mathbb{K} -Vektorraum wird.

II.6.3. SATZ. *Ist W ein Teilraum eines \mathbb{K} -Vektorraums V , dann bildet V/W mit obigen Operationen einen Vektorraum über \mathbb{K} . Die kanonische Abbildung $\pi: V \rightarrow V/W$, $\pi(v) := [v]$, ist eine lineare Surjektion mit $\ker(\pi) = W$. Zu jeder linearen Abbildung $\varphi: V \rightarrow U$ mit $\varphi|_W = 0$ existiert eine eindeutige lineare Abbildung $\bar{\varphi}: V/W \rightarrow U$, sodass $\varphi = \bar{\varphi} \circ \pi$.*

BEWEIS. Zunächst sind die Vektorraumaxiome (V1) – (V8) für V/W zu überprüfen. Wir beginnen mit der Assoziativität der Addition:

$$\begin{aligned} ([v_1] + [v_2]) + [v_3] &= [v_1 + v_2] + [v_3] = [(v_1 + v_2) + v_3] \\ &= [v_1 + (v_2 + v_3)] = [v_1] + [v_2 + v_3] = [v_1] + ([v_2] + [v_3]). \end{aligned}$$

Dies zeigt, dass die Addition auf V/W dem Axiom (V1) genügt. Analog erhalten wir aus Axiom (V2) für V sofort $[v] + [0] = [v + 0] = [v]$. Somit ist die Klasse $[0] \in V/W$ neutrales Element der Addition auf V/W , d.h. die Addition auf V/W erfüllt auch Axiom (V2). Ebenso erhalten wir $[v] + [-v] = [v + (-v)] = [0]$, d.h. die Klasse $[-v] \in V/W$ ist das additive Inverse der Klasse $[v]$. Die Addition auf V/W genügt daher auch Axiom (V3). Auch die Kommutativität der Addition auf V/W folgt sofort aus der Kommutativität der Addition in V , $[v_1] + [v_2] = [v_1 + v_2] = [v_2 + v_1] = [v_2] + [v_1]$, also genügt V/W Axiom (V4). Für $\lambda, \mu \in \mathbb{K}$ erhalten wir ähnlich $\lambda(\mu[v]) = \lambda[\mu v] = [\lambda(\mu v)] = [(\lambda\mu)v] = (\lambda\mu)[v]$, die Skalarmultiplikation in V/W genügt daher dem Axiom (V5). Weiters haben wir:

$$\begin{aligned} \lambda([v_1] + [v_2]) &= \lambda[v_1 + v_2] = [\lambda(v_1 + v_2)] \\ &= [\lambda v_1 + \lambda v_2] = [\lambda v_1] + [\lambda v_2] = \lambda[v_1] + \lambda[v_2]. \end{aligned}$$

In V/W gilt daher das Distributivgesetz (V6). Das andere Distributivgesetz (V7) lässt sich genauso verifizieren, $(\lambda + \mu)[v] = [(\lambda + \mu)v] = [\lambda v + \mu v] = [\lambda v] + [\mu v] = \lambda[v] + \mu[v]$, also genügt V/W auch Axiom (V7). Schließlich ist $1[v] = [1v] = [v]$ und damit ist auch Axiom (V8) erfüllt. Addition und Skalarmultiplikation auf V/W erfüllen somit alle Vektorraumaxiome und bilden daher einen \mathbb{K} -Vektorraum.

Die kanonische Abbildung $\pi: V \rightarrow V/W$ ist offensichtlich linear, denn es gilt $\pi(v + w) = [v + w] = [v] + [w] = \pi(v) + \pi(w)$ und auch $\pi(\lambda v) = [\lambda v] = \lambda[v] = \lambda\pi(v)$.³ Die Surjektivität von π ist trivial, jede Äquivalenzklasse besitzt ja mindestens einen Repräsentanten. Für den Kern von π folgt $\ker(\pi) = \pi^{-1}([0]) = \{v \in V \mid \pi(v) = [0]\} = \{v \in V \mid [v] = [0]\} = \{v \in V \mid v \sim 0\} = W$.

Sei nun $\varphi: V \rightarrow U$ eine lineare Abbildung die auf W verschwindet, d.h. $\varphi|_W = 0$. Für $v, v' \in V$ mit $v \sim v'$ gilt $v' - v \in W$, also $\varphi(v') - \varphi(v) = \varphi(v' - v) = 0$ und somit $\varphi(v) = \varphi(v')$. Daher ist die Abbildung

$$\bar{\varphi}: V/W \rightarrow U, \quad \bar{\varphi}([v]) := \varphi(v),$$

wohldefiniert. Auch ist sie offensichtlich linear, denn $\bar{\varphi}([v_1] + [v_2]) = \bar{\varphi}([v_1 + v_2]) = \varphi(v_1 + v_2) = \varphi(v_1) + \varphi(v_2) = \bar{\varphi}([v_1]) + \bar{\varphi}([v_2])$ und $\bar{\varphi}(\lambda[v]) = \bar{\varphi}([\lambda v]) = \varphi(\lambda v) = \lambda\varphi(v) = \lambda\bar{\varphi}([v])$. Schließlich gilt $(\bar{\varphi} \circ \pi)(v) = \bar{\varphi}(\pi(v)) = \bar{\varphi}([v]) = \varphi(v)$ und daher $\bar{\varphi} \circ \pi = \varphi$. Die Eindeutigkeit der Abbildung $\bar{\varphi}$ folgt sofort aus der Surjektivität von π . \square

³Die Vektorraumoperationen auf V/W wurden genau so definiert, dass π linear wird.

II.6.4. DEFINITION (Quotientenraum). Der Vektorraum V/W wird *Quotientenraum* V *nach* W oder *V modulo W* genannt. Die Abbildung $\pi: V \rightarrow V/W$, $\pi(v) = [v]$, wird als *kanonische Projektion* bezeichnet.

II.6.5. BEMERKUNG. Es gilt stets $V/\{0\} = V$ und $V/V = \{0\}$. Für einen Teilraum W eines Vektorraums V gilt $V/W = \{0\}$ genau dann, wenn $V = W$.

II.6.6. KOROLLAR. *Jede lineare Abbildung $\varphi: V \rightarrow W$ induziert einen linearen Isomorphismus*

$$\bar{\varphi}: V/\ker(\varphi) \xrightarrow{\cong} \text{img}(\varphi), \quad \bar{\varphi}([v]) = \varphi(v).$$

BEWEIS. Wir fassen φ als surjektive lineare Abbildung $\varphi: V \rightarrow \text{img}(\varphi)$ auf. Da φ auf $\ker(\varphi)$ verschwindet stellt $\bar{\varphi}: V/\ker(\varphi) \rightarrow \text{img}(\varphi)$, $\bar{\varphi}([v]) = \varphi(v)$, eine wohldefinierte lineare Abbildung dar, siehe Satz II.6.3. Offensichtlich ist $\bar{\varphi}$ surjektiv. Weiters gilt $\ker(\bar{\varphi}) = \{[0]\}$, denn aus $\bar{\varphi}([v]) = 0$ erhalten wir $\varphi(v) = 0$, also $v \in \ker(\varphi)$ und daher $[v] = [0] \in V/\ker(\varphi)$. Nach Proposition II.3.22 ist $\bar{\varphi}$ daher auch injektiv. Somit ist $\bar{\varphi}$ eine lineare Bijektion, also ein Isomorphismus, siehe Bemerkung II.3.11. \square

II.6.7. KOROLLAR. *Sei W ein Teilraum eines Vektorraums V und W' ein Komplement von W in V , d.h. $V = W \oplus W'$. Dann ist die Einschränkung der kanonischen Projektion, $\pi|_{W'}: W' \rightarrow V/W$, ein linearer Isomorphismus. Jede Äquivalenzklasse in V/W besitzt daher einen eindeutigen Repräsentanten in W' .*

BEWEIS. Dies folgt sofort aus Satz II.6.3 und Proposition II.5.13. \square

II.6.8. BEISPIEL. Sei $A \in M_{n \times n}(\mathbb{K})$ eine Matrix und $L := \{x \in \mathbb{K}^n : Ax = 0\}$ der Lösungsraum des assoziierten homogenen Gleichungssystems. Weiters sei $\xi \in \mathbb{K}^n$ und $y := A\xi \in \mathbb{K}^m$, also $A\xi = y$. Dann besteht die von ξ repräsentierte Äquivalenzklasse, $[\xi] \in \mathbb{K}^n/L$, genau aus jenen Vektoren, die das selbe Gleichungssystem wie ξ lösen, d.h. $[\xi] = \{x \in \mathbb{K}^n : Ax = y\}$.

II.6.9. BEISPIEL. Sei $[a, b]$ ein Intervall, und bezeichne $W \subseteq F([a, b], \mathbb{R})$ den Teilraum der konstanten Funktionen. Die von einer Funktion $f \in F([a, b], \mathbb{R})$ repräsentierte Äquivalenzklasse $[f] \in F([a, b], \mathbb{R})/W$ besteht daher aus allen Funktionen, die sich von f nur durch eine additive Konstante unterscheiden. Elemente von $F([a, b], \mathbb{R})/W$ können als Funktionen verstanden werden, die nur bis auf eine additive Konstante definiert sind.

II.6.10. BEISPIEL (Landau Symbol). Betrachte den Teilraum

$$W := o(1) = \{f \in F(\mathbb{R}, \mathbb{R}) \mid \lim_{x \rightarrow \infty} f(x) = 0\}$$

von $F(\mathbb{R}, \mathbb{R})$. Die von einer Funktion $f \in F(\mathbb{R}, \mathbb{R})$ repräsentierte Äquivalenzklasse $[f] \in F(\mathbb{R}, \mathbb{R})/W$ besteht daher genau aus jenen Funktionen g , für die $\lim_{x \rightarrow \infty} (g(x) - f(x)) = 0$ gilt, d.h. aus allen Funktionen die sich asymptotisch wie f verhalten. Elemente von $F(\mathbb{R}, \mathbb{R})/W$ können daher als Asymptoten verstanden werden.

III. Basen

Das Konzept der Basis bildet das wesentliche Hilfsmittel um die Struktur allgemeiner Vektorräume zu verstehen. Ist ein Vektorraum mit einer Basis ausgestattet, dann kann jeder Vektor in eindeutiger Weise als Linearkombination der Basisvektoren geschrieben werden, Vektoren können somit durch Zahlen, und zwar eine für jeden Basisvektor, beschrieben werden.

Basen werden als linear unabhängige Erzeugendensysteme definiert, wir beginnen dieses Kapitel daher mit zwei Abschnitten, in denen wir die Begriffe *Erzeugendensystem* und *lineare Unabhängigkeit* erläutern. In Abschnitt III.3 werden wir dann zeigen, dass jeder Vektorraum eine Basis besitzt und einige Konsequenzen dieses zentralen Resultats besprechen. Bevor wir im folgenden Kapitel IV näher auf die wichtige Klasse der endlich-dimensionalen Vektorräume eingehen, werden wir in Abschnitt III.4 noch grundlegende Eigenschaften des Dualraums zusammenstellen.

III.1. Erzeugendensysteme. Sei A eine Teilmenge eines \mathbb{K} -Vektorraums V . Unter einer *Linearkombination* von Elementen in A verstehen wir jeden Ausdruck der Form

$$\lambda_1 a_1 + \cdots + \lambda_n a_n,$$

wobei $\lambda_1, \dots, \lambda_n \in \mathbb{K}$ und $a_1, \dots, a_n \in A$. Auch $n = 0$ ist hier zulässig, wir interpretieren diesen Fall als leere Summe, ihr Wert ist definitionsgemäß gleich 0. Lässt sich ein Vektor $v \in V$ als Linearkombination von Elementen in A schreiben, d.h. existieren $a_i \in A$ und $\lambda_i \in \mathbb{K}$, sodass $v = \lambda_1 a_1 + \cdots + \lambda_n a_n$, dann können wir durch geeignetes Zusammenfassen auch erreichen, dass die Elemente a_1, \dots, a_n paarweise verschieden sind, siehe (V7). Somit lässt sich v auch in der Form

$$v = \sum_{a \in A} \lambda_a a$$

schreiben, wobei die Skalare $\lambda_a \in \mathbb{K}$ fast alle verschwinden, d.h. alle bis auf endlich viele gleich 0 sind.

III.1.1. DEFINITION (Lineare Hülle). Ist A eine Teilmenge eines \mathbb{K} -Vektorraums V , dann wird die Menge aller Vektoren, die sich als Linearkombinationen von Elementen in A schreiben lassen, d.h.

$$\langle A \rangle := \{ \lambda_1 a_1 + \cdots + \lambda_n a_n \mid \lambda_1, \dots, \lambda_n \in \mathbb{K}, a_1, \dots, a_n \in A \},$$

als *lineare Hülle* oder *lineares Erzeugnis* von A bezeichnet. Auch wird $\langle A \rangle$ der von A erzeugte bzw. *aufgespannte Teilraum* genannt. Sind $v_1, \dots, v_n \in V$, dann schreiben wir auch

$$\langle v_1, \dots, v_n \rangle := \langle \{v_1, \dots, v_n\} \rangle = \{ \lambda_1 v_1 + \cdots + \lambda_n v_n \mid \lambda_1, \dots, \lambda_n \in \mathbb{K} \}$$

für den von $\{v_1, \dots, v_n\}$ aufgespannten Teilraum.

III.1.2. PROPOSITION. Sei A eine Teilmenge eines Vektorraums V . Dann ist $\langle A \rangle$ der kleinste Teilraum von V , der A enthält, d.h. $\langle A \rangle$ bildet einen Teilraum von V , es gilt $A \subseteq \langle A \rangle$ und für jeden weiteren Teilraum W von V mit $A \subseteq W$ gilt schon $\langle A \rangle \subseteq W$. Insbesondere ist

$$\langle A \rangle = \bigcap_{\substack{W \text{ ist Teilraum von } V \\ \text{und } A \subseteq W}} W$$

wobei der Durchschnitt aller Teilräume von V gemeint ist, die A enthalten.

BEWEIS. Zunächst ist $\langle A \rangle$ nicht leer, denn $0 \in \langle A \rangle$. Weiters ist $\langle A \rangle$ offensichtlich abgeschlossen unter Addition. Um auch die Abgeschlossenheit unter Skalarmultiplikation zu sehen, sei nun $v \in \langle A \rangle$ und $\lambda \in \mathbb{K}$. Nach Definition der linearen Hülle lässt sich v in der Form $v = \lambda_1 a_1 + \dots + \lambda_n a_n$ schreiben, wobei $a_i \in A$ und $\lambda_i \in \mathbb{K}$. Es folgt $\lambda v = \lambda(\lambda_1 a_1 + \dots + \lambda_n a_n) = (\lambda \lambda_1) a_1 + \dots + (\lambda \lambda_n) a_n$, also ist auch $\lambda v \in \langle A \rangle$. Dies zeigt, dass $\langle A \rangle$ einen Teilraum von V bildet. Da sich jedes $a \in A$ als Linearkombination $a = 1a$ schreiben lässt, gilt auch $A \subseteq \langle A \rangle$. Sei nun W ein Teilraum von V , sodass $A \subseteq W$. Für beliebige $a_1, \dots, a_n \in A$ und $\lambda_1, \dots, \lambda_n \in \mathbb{K}$ folgt dann $\lambda_1 a_1 + \dots + \lambda_n a_n \in W$, also $\langle A \rangle \subseteq W$. \square

III.1.3. BEISPIEL. Etwa gilt in \mathbb{R}^3 :

$$\begin{aligned} \left\langle \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} \right\rangle &= \left\{ \lambda \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} : \lambda \in \mathbb{R} \right\} = \left\{ \begin{pmatrix} \lambda \\ 2\lambda \\ 3\lambda \end{pmatrix} : \lambda \in \mathbb{R} \right\} \\ \left\langle \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}, \begin{pmatrix} 3 \\ 2 \\ 1 \end{pmatrix} \right\rangle &= \left\{ \lambda \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} + \mu \begin{pmatrix} 3 \\ 2 \\ 1 \end{pmatrix} : \lambda, \mu \in \mathbb{R} \right\} = \left\{ \begin{pmatrix} \lambda+3\mu \\ 2\lambda+2\mu \\ 3\lambda+\mu \end{pmatrix} : \lambda, \mu \in \mathbb{R} \right\} \\ \left\langle \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}, \begin{pmatrix} 2 \\ 4 \\ 6 \end{pmatrix} \right\rangle &= \left\{ \lambda \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} + \mu \begin{pmatrix} 2 \\ 4 \\ 6 \end{pmatrix} : \lambda, \mu \in \mathbb{R} \right\} = \left\langle \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} \right\rangle \\ \left\langle \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right\rangle &= \left\{ \lambda \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + \mu \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} + \nu \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} : \lambda, \mu, \nu \in \mathbb{R} \right\} = \mathbb{R}^3 \end{aligned}$$

III.1.4. LEMMA. Sind A und B Teilmengen eines Vektorraums V dann gilt:

- (a) $\langle \emptyset \rangle = \{0\}$ und $\langle V \rangle = V$.
- (b) Ist $A \subseteq B$, dann auch $\langle A \rangle \subseteq \langle B \rangle$.
- (c) A ist genau dann Teilraum von V , wenn $A = \langle A \rangle$.
- (d) Es ist $B \subseteq \langle A \rangle$ genau dann, wenn $\langle A \cup B \rangle = \langle A \rangle$.
- (e) Es gilt $\langle A \cup B \rangle = \langle A \rangle + \langle B \rangle$.
- (f) Für jede lineare Abbildung $\varphi: V \rightarrow W$ gilt $\langle \varphi(A) \rangle = \varphi(\langle A \rangle)$.
- (g) Für je zwei lineare Abb. $\varphi, \psi: V \rightarrow W$ gilt: $\varphi|_A = \psi|_A \Leftrightarrow \varphi|_{\langle A \rangle} = \psi|_{\langle A \rangle}$.

BEWEIS. Die Behauptungen (a) und (b) sind trivial. Punkt (c) folgt sofort aus Proposition III.1.2. Ad (d): Gilt $\langle A \cup B \rangle = \langle A \rangle$, so folgt $B \subseteq A \cup B \subseteq \langle A \cup B \rangle = \langle A \rangle$, also $B \subseteq \langle A \rangle$. Für die andere Implikation sei nun $B \subseteq \langle A \rangle$. Da auch $A \subseteq \langle A \rangle$ erhalten wir $A \cup B \subseteq \langle A \rangle$ und somit $\langle A \cup B \rangle \subseteq \langle A \rangle$, denn $\langle A \cup B \rangle$ ist der kleinste Teilraum, der $A \cup B$ enthält. Aus $A \subseteq A \cup B$ folgt mit (b) aber auch die umgekehrte Inklusion, $\langle A \rangle \subseteq \langle A \cup B \rangle$, und somit Gleichheit.

Ad (e): Aus (b) erhalten wir zunächst $\langle A \rangle \cup \langle B \rangle \subseteq \langle A \cup B \rangle$, also $\langle A \rangle + \langle B \rangle \subseteq \langle A \cup B \rangle$, denn $\langle A \rangle + \langle B \rangle$ ist der kleinste Teilraum in dem $\langle A \rangle \cup \langle B \rangle$ enthalten ist, vgl. Proposition II.5.2. Andererseits ist $A \cup B \subseteq \langle A \rangle \cup \langle B \rangle \subseteq \langle A \rangle + \langle B \rangle$, es gilt daher auch die umgekehrte Inklusion, $\langle A \cup B \rangle \subseteq \langle A \rangle + \langle B \rangle$, denn $\langle A \cup B \rangle$ ist der kleinste Teilraum, der $A \cup B$ enthält.

Ad (f): Aus $A \subseteq \langle A \rangle$ folgt $\varphi(A) \subseteq \varphi(\langle A \rangle)$. Da $\varphi(\langle A \rangle)$ einen Teilraum bildet erhalten wir $\langle \varphi(A) \rangle \subseteq \varphi(\langle A \rangle)$, denn $\langle \varphi(A) \rangle$ ist der kleinste Teilraum, der $\varphi(A)$ enthält. Es gilt auch die umgekehrte Inklusion, $\varphi(\langle A \rangle) \subseteq \langle \varphi(A) \rangle$, denn jedes $v \in \langle A \rangle$ lässt sich als Linearkombination $v = \lambda_1 a_1 + \cdots + \lambda_n a_n$ schreiben, $a_i \in A$, $\lambda_i \in \mathbb{K}$, also $\varphi(v) = \varphi(\lambda_1 a_1 + \cdots + \lambda_n a_n) = \lambda_1 \varphi(a_1) + \cdots + \lambda_n \varphi(a_n) \in \langle \varphi(A) \rangle$.

Ad (g): Die eine Implikation ist trivial, da $A \subseteq \langle A \rangle$. Für die andere Implikation sei nun $\varphi|_A = \psi|_A$. Der Teilraum $V' := \{v \in V : \varphi(v) = \psi(v)\} = \ker(\varphi - \psi)$ enthält daher A als Teilmenge. Es folgt $\langle A \rangle \subseteq V'$ und daher $\varphi|_{\langle A \rangle} = \psi|_{\langle A \rangle}$. \square

III.1.5. BEMERKUNG. Aus Lemma III.1.4(b) folgt zwar $\langle A \cap B \rangle \subseteq \langle A \rangle \cap \langle B \rangle$, i.A. ist jedoch $\langle A \cap B \rangle \neq \langle A \rangle \cap \langle B \rangle$, siehe Übungsaufgabe 56.

III.1.6. DEFINITION (Erzeugendensystem). Eine Teilmenge E eines \mathbb{K} -Vektorraums V wird *Erzeugendensystem* von V genannt, wenn $\langle E \rangle = V$ gilt. In diesem Fall sagen wir auch E *erzeugt* V oder V wird von E *aufgespannt*. Ein Vektorraum heißt *endlich erzeugt*, wenn er ein endliches Erzeugendensystem besitzt.

III.1.7. BEMERKUNG. Jeder Vektorraum V besitzt ein Erzeugendensystem, etwa gilt $V = \langle V \rangle$ nach Lemma III.1.4(a).

III.1.8. BEMERKUNG. Ist E ein Erzeugendensystem eines Vektorraums V und gilt $E \subseteq E' \subseteq V$, dann ist auch E' Erzeugendensystem von V . Mit Lemma III.1.4(b) folgt nämlich $V = \langle E \rangle \subseteq \langle E' \rangle \subseteq V$, also $\langle E' \rangle = V$.

III.1.9. BEMERKUNG. Ist $\varphi: V \rightarrow W$ linear und E ein Erzeugendensystem von V , dann bildet $\varphi(E)$ ein Erzeugendensystem des Vektorraums $\text{img}(\varphi)$. Dies folgt sofort aus Lemma III.1.4(f). Insbesondere sind Quotienten endlich erzeugter Vektorräume wieder endlich erzeugt, vgl. Satz II.6.3.

Erzeugendensysteme lassen sich auch mit Hilfe linearer Abbildungen charakterisieren.

III.1.10. PROPOSITION (Erzeugendensysteme). *Für eine Teilmenge E eines \mathbb{K} -Vektorraums V sind die folgenden Aussagen äquivalent:*

- (a) E ist ein Erzeugendensystem von V .
- (b) Jedes $v \in V$ lässt sich als Linearkombination $v = \sum_{e \in E} \lambda_e e$ schreiben, für gewisse Skalare $\lambda_e \in \mathbb{K}$, die fast alle verschwinden.
- (c) Für je zwei lineare Abbildungen $\varphi, \psi: V \rightarrow W$ gilt: $\varphi|_E = \psi|_E \Rightarrow \varphi = \psi$.

BEWEIS. Die Äquivalenz (a) \Leftrightarrow (b) ist trivial. Die Implikation (a) \Rightarrow (c) folgt aus Lemma III.1.4(g). Ad (c) \Rightarrow (a): Betrachte die kanonische Projektion $\pi: V \rightarrow V/\langle E \rangle$, siehe Satz II.6.3. Da $\pi|_E = 0$, folgt aus der Eindeutigkeitsaussage in (c),

dass $\pi = 0$ gilt. Zusammen mit der Surjektivität von π erhalten wir $V/\langle E \rangle = \text{img}(\pi) = \text{img}(0) = \{0\}$ und somit $V = \langle E \rangle$, vgl. Bemerkung II.6.5. \square

III.1.11. BEISPIEL. Die Teilmenge $\left\{\begin{pmatrix} 1 \\ 2 \end{pmatrix}, \begin{pmatrix} 3 \\ 7 \end{pmatrix}\right\}$ bildet ein Erzeugendensystem von \mathbb{R}^2 , denn jedes Element von \mathbb{R}^2 lässt sich als Linearkombination dieser Vektoren schreiben,

$$\begin{pmatrix} x \\ y \end{pmatrix} = (7x - 3y) \begin{pmatrix} 1 \\ 2 \end{pmatrix} + (2x + y) \begin{pmatrix} 3 \\ 7 \end{pmatrix}.$$

III.1.12. BEISPIEL. Die Menge $\left\{\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 3 \\ 2 \\ 1 \end{pmatrix}\right\}$ bildet ein Erzeugendensystem von \mathbb{R}^3 , denn jedes Element von \mathbb{R}^3 lässt sich als Linearkombination dieser Vektoren schreiben,

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = (x - 2y + z) \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + (y - 2z) \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix} + z \begin{pmatrix} 3 \\ 2 \\ 1 \end{pmatrix}.$$

III.1.13. BEISPIEL. Die Menge $\left\{\begin{pmatrix} \mathbf{i} \\ 1 + \mathbf{i} \end{pmatrix}, \begin{pmatrix} 1 - \mathbf{i} \\ -3\mathbf{i} \end{pmatrix}\right\}$ bildet ein Erzeugendensystem von \mathbb{C}^2 , denn jedes Element von \mathbb{C}^2 lässt sich als Linearkombination dieser Vektoren schreiben,

$$\begin{pmatrix} z_1 \\ z_2 \end{pmatrix} = (-3\mathbf{i}z_1 - (1 - \mathbf{i})z_2) \begin{pmatrix} \mathbf{i} \\ 1 + \mathbf{i} \end{pmatrix} + (-(1 + \mathbf{i})z_1 + \mathbf{i}z_2) \begin{pmatrix} 1 - \mathbf{i} \\ -3\mathbf{i} \end{pmatrix}.$$

III.1.14. BEISPIEL. Der Vektorraum \mathbb{K}^n ist endlich erzeugt, die Menge der sogenannten *Einheitsvektoren*, $\{e_1, \dots, e_n\}$,

$$e_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad e_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad \dots, \quad e_n = \begin{pmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix},$$

bildet ein Erzeugendensystem von \mathbb{K}^n . Jedes $x \in \mathbb{K}^n$ lässt sich nämlich als Linearkombination $x = x_1e_1 + \dots + x_n e_n$ schreiben, wobei x_i die i -te Komponente von x bezeichnet.

III.1.15. BEISPIEL. Die Menge der Matrizen $\left\{\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}\right\}$ bildet ein Erzeugendensystem von $M_{2 \times 2}(\mathbb{K})$, denn jedes Element von $M_{2 \times 2}(\mathbb{K})$ lässt sich als Linearkombination dieser Matrizen schreiben,

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

III.1.16. BEISPIEL. Die Menge der Monome $\{1, z, z^2, z^3, \dots\}$ bildet ein Erzeugendensystem des Vektorraums der Polynome, $\mathbb{K}[z]$. Die Menge der Monome $\{1, z, z^2, \dots, z^n\}$ bildet ein Erzeugendensystem von $\mathbb{K}[z]_{\leq n}$.

Wir beenden diesen Abschnitt mit folgender Charakterisierung surjektiver linearer Abbildungen mittels Erzeugendensystemen.

III.1.17. PROPOSITION (Surjektive lineare Abbildungen). Für eine lineare Abbildung $\varphi: V \rightarrow W$ sind die folgenden Aussagen äquivalent:

- (a) φ ist surjektiv.
- (b) Für jedes Erzeugendensystem E von V ist $\varphi(E)$ Erzeugendensystem von W .
- (c) Es gibt eine Teilmenge $A \subseteq V$, sodass $\varphi(A)$ Erzeugendensystem von W ist.

BEWEIS. Die Implikation (a) \Rightarrow (b) folgt aus Bemerkung III.1.9. Die Implikation (b) \Rightarrow (c) ist offensichtlich, denn jeder Vektorraum besitzt ein Erzeugendensystem. Ad (c) \Rightarrow (a): Nach Voraussetzung existiert eine Teilmenge $A \subseteq V$, sodass $\langle \varphi(A) \rangle = W$. Aus Lemma III.1.4(f) folgt daher $\varphi(\langle A \rangle) = \langle \varphi(A) \rangle = W$, also ist φ surjektiv. \square

III.1.18. BEMERKUNG. Sei $A \in M_{m \times n}(\mathbb{K})$ eine Matrix und $\psi: \mathbb{K}^n \rightarrow \mathbb{K}^m$, $\psi(x) := Ax$, die damit assoziierte lineare Abbildung. Dann bildet die Menge der Spaltenvektoren von A , d.h.

$$\left\{ \begin{pmatrix} a_{11} \\ \vdots \\ a_{m1} \end{pmatrix}, \dots, \begin{pmatrix} a_{1n} \\ \vdots \\ a_{mn} \end{pmatrix} \right\} = \{\psi(e_1), \dots, \psi(e_n)\} \subseteq \mathbb{K}^m$$

ein Erzeugendensystem von $\text{img}(\psi)$, wobei e_1, \dots, e_n die Einheitsvektoren in \mathbb{K}^n bezeichnen. Dies folgt aus Bemerkung III.1.9, denn $E = \{e_1, \dots, e_n\}$ bildet ein Erzeugendensystem von \mathbb{K}^n . Der Teilraum jener $y \in \mathbb{K}^m$, für die das Gleichungssystem $Ax = y$ mindestens eine Lösung $x \in \mathbb{K}^n$ besitzt, wird daher von den Spalten von A aufgespannt, vgl. Bemerkung II.4.13. Insbesondere ist ψ genau dann surjektiv, wenn die Spalten von A ein Erzeugendensystem von \mathbb{K}^m bilden. Etwa ist die mit der Matrix

$$A = \begin{pmatrix} 2 & 0 & 1 & 1 & 9 & 2 & 0 \\ 3 & 1 & 4 & 0 & 8 & 4 & 1 \\ 5 & 1 & 7 & 0 & 3 & 6 & 0 \end{pmatrix}$$

assoziierte lineare Abbildung $\psi: \mathbb{K}^7 \rightarrow \mathbb{K}^3$, $\psi(x) := Ax$, surjektiv, denn die Menge der Spaltenvektoren von A enthält die Vektoren e_1 , e_2 und $e_2 + e_3$, deren lineare Hülle enthält daher auch den Vektor $e_3 = (e_2 + e_3) - e_2$ und stimmt somit mit \mathbb{K}^3 überein, vgl. Beispiel III.1.14. Das lineare Gleichungssystem $Ax = y$ besitzt daher für jedes $y \in \mathbb{K}^3$ mindestens eine Lösung $x \in \mathbb{K}^7$.

III.2. Lineare Unabhängigkeit. Eine Teilmenge eines Vektorraums heißt linear abhängig wenn zwischen ihren Elementen nicht-triviale lineare Relationen bestehen. Genauer haben wir folgende Definition.

III.2.1. DEFINITION (Lineare Abhängigkeit). Es sei V ein Vektorraum über \mathbb{K} . Eine Teilmenge $A \subseteq V$ wird *linear abhängig* genannt, falls gilt: Es existieren paarweise verschiedene Elemente $a_1, \dots, a_n \in A$ und Skalare $\lambda_1, \dots, \lambda_n \in \mathbb{K}$, die nicht alle verschwinden, sodass $\lambda_1 a_1 + \dots + \lambda_n a_n = 0$.

III.2.2. BEISPIEL. Die Teilmenge $\left\{\begin{pmatrix} 1 \\ 2 \end{pmatrix}, \begin{pmatrix} 3 \\ 7 \end{pmatrix}, \begin{pmatrix} 2 \\ 4 \end{pmatrix}\right\}$ von \mathbb{R}^2 ist linear abhängig, denn $2\begin{pmatrix} 1 \\ 2 \end{pmatrix} + (-1)\begin{pmatrix} 2 \\ 4 \end{pmatrix} = 0$.

III.2.3. BEISPIEL. Die Teilmenge $\left\{\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}, \begin{pmatrix} 2 \\ 3 \\ 4 \end{pmatrix}, \begin{pmatrix} 3 \\ 5 \\ 7 \end{pmatrix}\right\}$ von \mathbb{R}^3 ist linear abhängig, denn $\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} + \begin{pmatrix} 2 \\ 3 \\ 4 \end{pmatrix} - \begin{pmatrix} 3 \\ 5 \\ 7 \end{pmatrix} = 0$.

III.2.4. BEMERKUNG. Jede Teilmenge eines Vektorraums, die 0 enthält ist linear abhängig, denn $0 = 1 \cdot 0$.

III.2.5. BEMERKUNG. Jede Obermenge einer linear abhängigen Menge ist linear abhängig, d.h. ist A eine linear abhängige Teilmenge eines Vektorraums V und gilt $A \subseteq A' \subseteq V$, dann ist auch A' linear abhängig.

III.2.6. PROPOSITION (Lineare Abhängigkeit). *Für eine Teilmenge A eines \mathbb{K} -Vektorraums V sind die folgenden Aussagen äquivalent:*

- (a) A ist linear abhängig in V .
- (b) Es existieren $\lambda_1, \dots, \lambda_n \in \mathbb{K}$ und paarweise verschiedene $a, a_1, \dots, a_n \in A$, sodass $a = \lambda_1 a_1 + \dots + \lambda_n a_n$, d.h. wenigstens eines der Elemente von A lässt sich als Linearkombination der restlichen schreiben.
- (c) Es existiert eine echte Teilmenge $A' \subsetneq A$, sodass $\langle A' \rangle = \langle A \rangle$.

BEWEIS. Ad (a) \Rightarrow (b): Nach Voraussetzung existieren $\lambda_1, \dots, \lambda_n \in \mathbb{K}$, nicht alle gleich 0, und paarweise verschiedene $a_1, \dots, a_n \in A$, sodass $\lambda_1 a_1 + \dots + \lambda_n a_n = 0$. Durch Umm Nummerieren dürfen wir o.B.d.A. $\lambda_n \neq 0$ annehmen. Es gilt daher

$$a_n = (-\lambda_n^{-1} \lambda_1) a_1 + \dots + (-\lambda_n^{-1} \lambda_{n-1}) a_{n-1},$$

die gewünschte Darstellung.

Für die Implikation (b) \Rightarrow (c) seien nun $\lambda_1, \dots, \lambda_n \in \mathbb{K}$ und $a, a_1, \dots, a_n \in A$ paarweise verschieden, sodass $a = \lambda_1 a_1 + \dots + \lambda_n a_n$. Dann ist $A' := A \setminus \{a\} \subsetneq A$ eine echte Teilmenge von A und es gilt $a_1, \dots, a_n \in A'$. Aus $a = \lambda_1 a_1 + \dots + \lambda_n a_n$ erhalten wir weiters $a \in \langle A' \rangle$. Nach Lemma III.1.4(d) gilt daher $\langle A' \cup \{a\} \rangle = \langle A' \rangle$. Da $A = A' \cup \{a\}$ folgt $\langle A \rangle = \langle A' \rangle$.

Ad (c) \Rightarrow (a): Sei also $A' \subsetneq A$ eine echte Teilmenge mit $\langle A' \rangle = \langle A \rangle$. Es existiert daher $a \in A \setminus A'$. Da $a \in A \subseteq \langle A \rangle = \langle A' \rangle$ existieren paarweise verschiedene $a_1, \dots, a_n \in A'$ und $\lambda_1, \dots, \lambda_n \in \mathbb{K}$ mit $a = \lambda_1 a_1 + \dots + \lambda_n a_n$. Auch die Vektoren a, a_1, \dots, a_n sind paarweise verschieden, denn $a \notin A'$ aber $a_1, \dots, a_n \in A'$. Da

$$(-1)a + \lambda_1 a_1 + \dots + \lambda_n a_n = 0,$$

ist A also linear abhängig. □

III.2.7. BEISPIEL. Wir wollen die Aussage von Proposition III.2.6 an der linear abhängigen Teilmenge $\left\{\begin{pmatrix} 1 \\ 2 \end{pmatrix}, \begin{pmatrix} 3 \\ 7 \end{pmatrix}, \begin{pmatrix} 2 \\ 4 \end{pmatrix}\right\} \subseteq \mathbb{R}^2$ aus Beispiel III.2.2 illustrieren. In diesem Fall lassen sich zwei Vektoren als Linearkombinationen der anderen schreiben, $\begin{pmatrix} 1 \\ 2 \end{pmatrix} = \frac{1}{2} \begin{pmatrix} 2 \\ 4 \end{pmatrix}$ und $\begin{pmatrix} 2 \\ 4 \end{pmatrix} = 2 \begin{pmatrix} 1 \\ 2 \end{pmatrix}$, aber der dritte Vektor ist nicht Linearkombination der restlichen, denn aus $\begin{pmatrix} 3 \\ 7 \end{pmatrix} = \lambda \begin{pmatrix} 1 \\ 2 \end{pmatrix} + \mu \begin{pmatrix} 2 \\ 4 \end{pmatrix}$ folgt $3 = \lambda + 2\mu$

und $7 = 2\lambda + 4\mu$ und dieses Gleichungssystem besitzt keine Lösungen. Auch gilt $\mathbb{R}^2 = \langle (\frac{1}{2}), (\frac{3}{7}) \rangle = \langle (\frac{3}{7}), (\frac{2}{4}) \rangle = \langle (\frac{1}{2}), (\frac{3}{7}), (\frac{2}{4}) \rangle \neq \langle (\frac{1}{2}), (\frac{2}{4}) \rangle = \langle (\frac{1}{2}) \rangle$.

III.2.8. BEISPIEL. Die Menge der Vektoren $\left\{ \begin{pmatrix} 0 \\ 1 \\ \mathbf{i} \end{pmatrix}, \begin{pmatrix} 1 \\ 2\mathbf{i} \\ 3+5\mathbf{i} \end{pmatrix}, \begin{pmatrix} \mathbf{i} \\ -3\mathbf{i} \\ -2+5\mathbf{i} \end{pmatrix} \right\}$ ist linear abhängig in \mathbb{C}^3 , denn

$$\begin{pmatrix} 1 \\ 2\mathbf{i} \\ 3+5\mathbf{i} \end{pmatrix} = (3+2\mathbf{i}) \begin{pmatrix} 0 \\ 1 \\ \mathbf{i} \end{pmatrix} - \mathbf{i} \begin{pmatrix} \mathbf{i} \\ -3\mathbf{i} \\ -2+5\mathbf{i} \end{pmatrix}.$$

III.2.9. BEISPIEL. Sei $\omega \in \mathbb{R} \setminus \frac{\pi}{2}\mathbb{Z}$ fix. Die Funktionen $f_1, f_2, f_3: \mathbb{R} \rightarrow \mathbb{R}$,

$$f_1(x) := \sin(x), \quad f_2(x) := \cos(x) \quad \text{und} \quad f_3(x) := \sin(x + \omega).$$

sind paarweise verschieden, denn $f_1(0) = 0 \neq f_3(0) \neq 1 = f_2(0)$. Die Teilmenge $\{f_1, f_2, f_3\} \subseteq F(\mathbb{R}, \mathbb{R})$ ist linear abhängig, denn nach dem Additionstheorem für Winkelfunktionen gilt $f_3 = \cos(\omega)f_1 + \sin(\omega)f_2$.

III.2.10. DEFINITION (Lineare Unabhängigkeit). Eine Teilmenge eines Vektorraums wird *linear unabhängig* genannt, wenn sie nicht linear abhängig ist.

III.2.11. PROPOSITION (Lineare Unabhängigkeit). Für eine Teilmenge A eines \mathbb{K} -Vektorraums V sind die folgenden Aussagen äquivalent:

- (a) A ist linear unabhängig in V .
- (b) Sind $a_1, \dots, a_n \in A$ paarweise verschieden und $\lambda_1, \dots, \lambda_n \in \mathbb{K}$, sodass $\lambda_1 a_1 + \dots + \lambda_n a_n = 0$, dann gilt schon $\lambda_1 = \dots = \lambda_n = 0$.
- (c) Ist $\sum_{a \in A} \lambda_a a = 0$, wobei die Skalare $\lambda_a \in \mathbb{K}$ fast alle verschwinden, dann gilt schon $\lambda_a = 0$ für alle $a \in A$.
- (d) Ist $\sum_{a \in A} \lambda_a a = \sum_{a \in A} \mu_a a$ wobei die Skalare $\lambda_a \in \mathbb{K}$ und $\mu_a \in \mathbb{K}$ fast alle verschwinden, dann gilt schon $\lambda_a = \mu_a$ für alle $a \in A$.
- (e) Für jede echte Teilmenge $A' \subsetneq A$ gilt $\langle A' \rangle \neq \langle A \rangle$.

BEWEIS. Die Äquivalenz (a) \Leftrightarrow (b) folgt sofort aus der Definition. Die Äquivalenz (b) \Leftrightarrow (c) ist trivial. Ad (c) \Rightarrow (d): Sei also $\sum_{a \in A} \lambda_a a = \sum_{a \in A} \mu_a a$, wobei die Skalare $\lambda_a \in \mathbb{K}$ und $\mu_a \in \mathbb{K}$ fast alle verschwinden. Es folgt

$$\sum_{a \in A} (\lambda_a - \mu_a) a = \sum_{a \in A} \lambda_a a - \sum_{a \in A} \mu_a a = 0.$$

Für jedes $a \in A$ gilt daher nach Voraussetzung $\lambda_a - \mu_a = 0$, also $\lambda_a = \mu_a$. Die Implikation (d) \Rightarrow (c) folgt sofort indem wir $\mu_a = 0$ setzen. Die Äquivalenz (a) \Leftrightarrow (e) folgt aus Proposition III.2.6. \square

III.2.12. BEMERKUNG. Die leere Menge ist stets linear unabhängig.

III.2.13. BEMERKUNG. Jede Teilmenge einer linear unabhängigen Menge ist linear unabhängig, d.h. ist A eine linear unabhängige Teilmenge eines Vektorraums V und gilt $A' \subseteq A$, dann ist auch A' linear unabhängig. Dies folgt aus Bemerkung III.2.5.

III.2.14. BEMERKUNG. Ist W ein Teilraum eines Vektorraums V und $A \subseteq W$ linear unabhängig in W , dann ist A auch linear unabhängig in V .

III.2.15. BEISPIEL. Eine 1-elementige Teilmenge $\{v\} \subseteq V$ ist genau dann linear unabhängig, wenn $v \neq 0$.

III.2.16. BEISPIEL. Die Menge der Vektoren $\left\{\begin{pmatrix} 1 \\ 2 \end{pmatrix}, \begin{pmatrix} 3 \\ 7 \end{pmatrix}\right\}$ ist linear unabhängig in \mathbb{R}^2 . Sind nämlich $\lambda, \mu \in \mathbb{R}$ und

$$\lambda \begin{pmatrix} 1 \\ 2 \end{pmatrix} + \mu \begin{pmatrix} 3 \\ 7 \end{pmatrix} = 0,$$

dann folgt $\lambda + 3\mu = 0$ und $2\lambda + 7\mu = 0$, also $\lambda = \mu = 0$.

III.2.17. BEISPIEL. Die Menge der Vektoren $\left\{\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 3 \\ 2 \\ 1 \end{pmatrix}\right\}$ ist linear unabhängig in \mathbb{R}^3 . Sind nämlich $\lambda, \mu, \nu \in \mathbb{R}$ und

$$\lambda \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + \mu \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix} + \nu \begin{pmatrix} 3 \\ 2 \\ 1 \end{pmatrix} = 0,$$

so folgt $\lambda + 2\mu + 3\nu = 0$, $\mu + 2\nu = 0$ und $\nu = 0$, also $\lambda = \mu = \nu = 0$.

III.2.18. BEISPIEL. Die Menge der Vektoren $\left\{\begin{pmatrix} \mathbf{i} \\ 1+\mathbf{i} \end{pmatrix}, \begin{pmatrix} 1-\mathbf{i} \\ -3\mathbf{i} \end{pmatrix}\right\}$ ist linear unabhängig in \mathbb{C}^2 . Sind nämlich $\lambda, \nu \in \mathbb{C}$ und

$$\lambda \begin{pmatrix} \mathbf{i} \\ 1+\mathbf{i} \end{pmatrix} + \mu \begin{pmatrix} 1-\mathbf{i} \\ -3\mathbf{i} \end{pmatrix} = 0,$$

dann folgt $\mathbf{i}\lambda + (1-\mathbf{i})\mu = 0$ und $(1+\mathbf{i})\lambda - 3\mathbf{i}\mu = 0$, also $\lambda = \mu = 0$.

III.2.19. BEISPIEL. Die Menge der Einheitsvektoren $\{e_1, \dots, e_n\}$ ist linear unabhängig in \mathbb{K}^n , denn aus $\lambda_1 e_1 + \dots + \lambda_n e_n = 0$, folgt $\lambda_1 = \dots = \lambda_n = 0$.

III.2.20. BEISPIEL. Die Menge der Matrizen $\left\{\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}\right\}$ ist linear unabhängig in $M_{2 \times 2}(\mathbb{K})$. Sind nämlich $\lambda_1, \lambda_2, \lambda_3, \lambda_4 \in \mathbb{K}$ und

$$\lambda_1 \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + \lambda_2 \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} + \lambda_3 \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} + \lambda_4 \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} = 0,$$

dann gilt schon $\lambda_1 = \lambda_2 = \lambda_3 = \lambda_4 = 0$.

III.2.21. BEISPIEL. Die Menge der Monome, $\{1, z, z^2, z^3, \dots\}$, ist linear unabhängig in $\mathbb{K}[z]$.

III.2.22. BEISPIEL. Die Teilmenge $\{\sin, \cos\} \subseteq F(\mathbb{R}, \mathbb{R})$ ist linear unabhängig. Sind nämlich $\lambda, \mu \in \mathbb{R}$ mit

$$\lambda \sin + \mu \cos = 0,$$

so folgt durch Auswerten bei 0 und $\pi/2$ sofort $0 = \lambda \sin(0) + \mu \cos(0) = \mu$ und $0 = \lambda \sin(\pi/2) + \mu \cos(\pi/2) = \lambda$.

III.2.23. BEISPIEL. Betrachte die Funktionen $f_1, f_2, f_3: \mathbb{R} \rightarrow \mathbb{R}$, $f_1(x) := e^x$, $f_2(x) := e^{2x}$ und $f_3(x) := e^{3x}$. Die Teilmenge $\{f_1, f_2, f_3\} \subseteq F(\mathbb{R}, \mathbb{R})$ ist linear unabhängig. Sind nämlich $\lambda, \mu, \nu \in \mathbb{R}$ und

$$\lambda f_1 + \mu f_2 + \nu f_3 = 0,$$

dann erhalten wir durch Auswerten bei $x = \ln 1$, $x = \ln 2$ und $x = \ln 3$ das lineare Gleichungssystem

$$\begin{aligned}\lambda + \nu + \mu &= 0 \\ 2\lambda + 4\nu + 8\mu &= 0 \\ 3\lambda + 9\nu + 27\mu &= 0\end{aligned}$$

und dieses besitzt nur eine Lösung, $\lambda = \mu = \nu = 0$.

III.3. Basen. Um Elemente eines Vektorraums effizient durch Linearkombinationen beschreiben zu können, liegt es nahe möglichst kleine, d.h. linear unabhängige Erzeugendensysteme zu betrachten. Diese werden Basen genannt.

III.3.1. DEFINITION (Basis). Unter einer *Basis* eines Vektorraums verstehen wir ein linear unabhängiges Erzeugendensystem.

III.3.2. PROPOSITION (Basen). *Für eine Teilmenge B eines \mathbb{K} -Vektorraums V , sind die folgenden Aussagen äquivalent:*

- (a) B ist eine Basis von V .
- (b) Jedes $v \in V$ lässt sich in eindeutiger Weise als Linearkombination $v = \sum_{b \in B} \lambda_b b$ schreiben, wobei die Skalare $\lambda_b \in \mathbb{K}$ fast alle verschwinden.
- (c) Zu jedem Vektorraum W und jeder Abbildung $f: B \rightarrow W$ existiert eine eindeutige lineare Abbildung $\varphi: V \rightarrow W$ mit $\varphi|_B = f$.

BEWEIS. Die Äquivalenz (a) \Leftrightarrow (b) folgt aus Proposition III.1.10(b) und Proposition III.2.11(d). Ad (b) \Rightarrow (c): Wir definieren $\varphi: V \rightarrow W$ durch $\varphi(v) := \sum_{b \in B} \lambda_b f(b)$, wobei $\lambda_b \in \mathbb{K}$ jene eindeutig bestimmten Skalare bezeichnen, für die $v = \sum_{b \in B} \lambda_b b$ gilt. Diese Abbildung ist wohldefiniert und es gilt $\varphi|_B = f$, denn jedes $b \in B$ lässt sich in der Form $b = 1 \cdot b$ schreiben, also $\varphi(b) = 1f(b) = f(b)$. Es bleibt daher nur noch die Linearität von φ zu zeigen. Seien dazu $v = \sum_{b \in B} \lambda_b b$ und $v' = \sum_{b \in B} \lambda'_b b$. Dann ist $v + v' = \sum_{b \in B} \lambda_b b + \sum_{b \in B} \lambda'_b b = \sum_{b \in B} (\lambda_b + \lambda'_b) b$ und daher

$$\varphi(v + v') = \sum_{b \in B} (\lambda_b + \lambda'_b) f(b) = \sum_{b \in B} \lambda_b f(b) + \sum_{b \in B} \lambda'_b f(b) = \varphi(v) + \varphi(v').$$

Für $\mu \in \mathbb{K}$ gilt analog $\mu v = \mu \sum_{b \in B} \lambda_b b = \sum_{b \in B} (\mu \lambda_b) b$ und daher $\varphi(\mu v) = \sum_{b \in B} (\mu \lambda_b) f(b) = \mu \sum_{b \in B} \lambda_b f(b) = \mu \varphi(v)$. Damit ist die Existenz einer linearen Abbildung φ mit den gewünschten Eigenschaften gezeigt. Die Eindeutigkeit von φ folgt aus Proposition III.1.10.

Ad (c) \Rightarrow (a): Betrachte die kanonische Projektion $\pi: V \rightarrow V/\langle B \rangle$. Da $\pi|_B = 0$, folgt aus der Eindeutigkeitsaussage in (c), dass $\pi = 0$ gilt. Da π surjektiv

ist, erhalten wir $V/\langle B \rangle = \text{img}(\pi) = \text{img}(0) = \{0\}$ und somit $V = \langle B \rangle$, vgl. Bemerkung II.6.5. Dies zeigt, dass B ein Erzeugendensystem von V bildet. Um auch die lineare Unabhängigkeit von B zu zeigen, seien $b_1, \dots, b_n \in B$ paarweise verschieden und $\lambda_1, \dots, \lambda_n \in \mathbb{K}$ so, dass $\lambda_1 b_1 + \dots + \lambda_n b_n = 0$. Zu jedem $i \in \{1, \dots, n\}$ existiert nach Voraussetzung eine lineare Abbildung $\varphi: V \rightarrow \mathbb{K}$ mit $\varphi(b_i) = 1$ und $\varphi|_{B \setminus \{b_i\}} = 0$. Es folgt

$$0 = \varphi(0) = \varphi(\lambda_1 b_1 + \dots + \lambda_n b_n) = \lambda_1 \varphi(b_1) + \dots + \lambda_n \varphi(b_n) = \lambda_i \varphi(b_i) = \lambda_i,$$

und daher $\lambda_1 = \dots = \lambda_n = 0$. Dies zeigt, dass B auch linear unabhängig ist. \square

III.3.3. BEMERKUNG. Ist A eine linear unabhängige Teilmenge eines Vektorraums V , dann bildet A eine Basis des Teilraums $\langle A \rangle$.

III.3.4. BEMERKUNG. Die leere Menge bildet eine Basis des trivialen Vektorraums $\{0\}$.

III.3.5. BEISPIEL. Die Menge der Einheitsvektoren $\{e_1, \dots, e_n\}$ bildet eine Basis von \mathbb{K}^n , siehe Beispiel III.1.14 und Beispiel III.2.19. Diese Basis wird als *Standardbasis* von \mathbb{K}^n bezeichnet.

III.3.6. BEISPIEL. Die Teilmenge $\{(\frac{1}{2}), (\frac{3}{7})\}$ ist eine Basis von \mathbb{R}^2 , denn nach Beispiel III.1.11 bildet sie ein Erzeugendensystem, das nach Beispiel III.2.16 auch linear unabhängig ist.

III.3.7. BEISPIEL. Die Menge $\{(\frac{1}{0}), (\frac{3}{2}), (\frac{5}{3})\}$ ist eine Basis von \mathbb{R}^3 , denn nach Beispiel III.1.12 bildet sie ein Erzeugendensystem, das nach Beispiel III.2.17 auch linear unabhängig ist.

III.3.8. BEISPIEL. Die Menge $\{(\frac{i}{1+i}), (\frac{1-i}{-3i})\}$ ist eine Basis von \mathbb{C}^2 , denn nach Beispiel III.1.13 bildet sie ein Erzeugendensystem, das nach Beispiel III.2.18 auch linear unabhängig ist.

III.3.9. BEISPIEL. Die Menge der Matrizen $\{(\frac{1}{0}), (\frac{0}{1}), (\frac{0}{1}), (\frac{0}{0})\}$ ist eine Basis von $M_{2 \times 2}(\mathbb{K})$, denn nach Beispiel III.1.15 bildet sie ein Erzeugendensystem, das nach Beispiel III.2.20 auch linear unabhängig ist.

III.3.10. BEISPIEL. Die Menge der Monome $\{1, z, z^2, z^3, \dots\}$ bildet eine Basis von $\mathbb{K}[z]$. Die Menge der Monome $\{1, z, z^2, \dots, z^n\}$ bildet eine Basis von $\mathbb{K}[z]_{\leq n}$.

III.3.11. BEMERKUNG. Zwei \mathbb{K} -Vektorräume, die gleichmächtige Basen besitzen, sind isomorph. Seien dazu V und W zwei \mathbb{K} -Vektorräume mit Basen B bzw. C , und sei $f: B \rightarrow C$ eine Bijektion. Nach Proposition III.3.2 existieren lineare Abbildungen $\varphi: V \rightarrow W$ und $\psi: W \rightarrow V$, sodass $\varphi|_B = f$ und $\psi|_C = f^{-1}$. Es folgt $(\psi \circ \varphi)|_B = f^{-1} \circ f = \text{id}_B = \text{id}_V|_B$ und $(\varphi \circ \psi)|_C = f \circ f^{-1} = \text{id}_C = \text{id}_W|_C$. Aus der Eindeutigkeitsaussage in Proposition III.3.2(c) erhalten wir also $\psi \circ \varphi = \text{id}_V$ und $\varphi \circ \psi = \text{id}_W$. Somit ist φ ein linearer Isomorphismus mit Umkehrabbildung $\varphi^{-1} = \psi$, d.h. $V \cong W$. Ist etwa B eine Basis eines \mathbb{K} -Vektorraums V , die aus

genau n Elementen besteht, dann folgt $V \cong \mathbb{K}^n$. Daraus erhalten wir erneut $M_{m \times n}(\mathbb{K}) \cong \mathbb{K}^{mn}$ und $\mathbb{K}[z]_{\leq n} \cong \mathbb{K}^{n+1}$.

III.3.12. PROPOSITION (Injektive lineare Abbildungen). *Ist $\varphi: V \rightarrow W$ linear und B eine Basis von V , dann sind die folgenden Aussagen äquivalent:*

- (a) φ ist injektiv.
- (b) φ bildet jede linear unabhängige Teilmenge von V bijektiv auf eine linear unabhängige Teilmenge von W ab.
- (c) φ bildet jede Basis von V bijektiv auf eine lin. unabh. Teilmenge von W ab.
- (d) φ bildet B bijektiv auf eine linear unabhängige Teilmenge von W ab.

BEWEIS. Ad (a) \Rightarrow (b): Sei also $A \subseteq V$ linear unabhängig. Da φ injektiv ist, wird A durch φ bijektiv auf $\varphi(A)$ abgebildet. Es genügt daher zu zeigen, dass $\varphi(A)$ linear unabhängig ist. Seien dazu $w_1, \dots, w_n \in \varphi(A)$ paarweise verschieden und $\lambda_1, \dots, \lambda_n \in \mathbb{K}$, sodass $\lambda_1 w_1 + \dots + \lambda_n w_n = 0$. Da $w_i \in \varphi(A)$, existieren $v_1, \dots, v_n \in A$, sodass $\varphi(v_i) = w_i$, für alle $i = 1, \dots, n$. Beachte, dass auch die Vektoren v_1, \dots, v_n paarweise verschieden sein müssen. Wir erhalten

$$\varphi(\lambda_1 v_1 + \dots + \lambda_n v_n) = \lambda_1 \varphi(v_1) + \dots + \lambda_n \varphi(v_n) = \lambda_1 w_1 + \dots + \lambda_n w_n = 0.$$

Aus der Injektivität von φ folgt somit $\lambda_1 v_1 + \dots + \lambda_n v_n = 0$. Da A linear unabhängig ist, erhalten wir schließlich $\lambda_1 = \dots = \lambda_n = 0$. Dies zeigt, dass die Menge $\varphi(A)$ linear unabhängig in W ist.

Die Implikationen (b) \Rightarrow (c) \Rightarrow (d) sind trivial.

Ad (d) \Rightarrow (a): Sei $v \in \ker(\varphi)$. Da B ein Erzeugendensystem von V bildet, existieren paarweise verschiedene $b_1, \dots, b_n \in B$ und $\lambda_1, \dots, \lambda_n \in \mathbb{K}$, sodass $v = \lambda_1 b_1 + \dots + \lambda_n b_n$. Es folgt $0 = \varphi(v) = \varphi(\lambda_1 b_1 + \dots + \lambda_n b_n) = \lambda_1 \varphi(b_1) + \dots + \lambda_n \varphi(b_n)$. Da $\varphi|_B$ injektiv ist, sind die Vektoren $\varphi(b_1), \dots, \varphi(b_n)$ paarweise verschieden. Da $\varphi(B)$ linear unabhängig ist, folgt $\lambda_1 = \dots = \lambda_n = 0$, also $v = 0$. Dies zeigt $\ker(\varphi) = \{0\}$, folglich ist φ injektiv, vgl. Proposition II.3.22. \square

III.3.13. BEMERKUNG. Sei $A \in M_{m \times n}(\mathbb{K})$ eine Matrix. Die mit A assoziierte lineare Abbildung $\psi: \mathbb{K}^n \rightarrow \mathbb{K}^m$, $\psi(x) = Ax$, ist genau dann injektiv, wenn die Spaltenvektoren von A paarweise verschieden sind und eine linear unabhängige Teilmenge von \mathbb{K}^m bilden. Dies folgt aus Proposition III.3.12, denn die Standardbasis $E = \{e_1, \dots, e_n\}$ von \mathbb{K}^n wird durch ψ auf die Menge der Spaltenvektoren von A abgebildet.

III.3.14. PROPOSITION (Isomorphismen). *Ist $\varphi: V \rightarrow W$ linear und B eine Basis von V , dann sind die folgenden Aussagen äquivalent:*

- (a) φ ist ein Isomorphismus.
- (b) φ bildet jede Basis von V bijektiv auf eine Basis von W ab.
- (c) φ bildet B bijektiv auf eine Basis von W ab.

BEWEIS. Die Implikation (a) \Rightarrow (b) folgt aus Proposition III.1.17 und Proposition III.3.12. Die Implikation (b) \Rightarrow (c) ist trivial. Die Implikation (c) \Rightarrow (a) folgt aus Proposition III.1.17 und Proposition III.3.12. \square

III.3.15. BEMERKUNG. Sei $A \in M_{m \times n}(\mathbb{K})$ eine Matrix. Die mit A assoziierte lineare Abbildung $\psi: \mathbb{K}^n \rightarrow \mathbb{K}^m$, $\psi(x) = Ax$, ist genau dann ein Isomorphismus, wenn die Spaltenvektoren von A paarweise verschieden sind und eine Basis von \mathbb{K}^m bilden. Dies folgt aus Proposition III.3.14, denn die Standardbasis $E = \{e_1, \dots, e_n\}$ von \mathbb{K}^n wird durch ψ auf die Menge der Spaltenvektoren von A abgebildet. Wir werden später sehen, dass in diesem Fall $m = n$ gelten muss.

III.3.16. PROPOSITION. *Seien W' und W'' zwei komplementäre Teilräume eines Vektorraums V , d.h. $V = W' \oplus W''$. Weiters sei B' eine Basis von W' und B'' eine Basis von W'' . Dann ist $B' \cap B'' = \emptyset$ und $B' \cup B''$ ist eine Basis von V .*

BEWEIS. Nach Lemma III.1.4(e) gilt $\langle B' \cup B'' \rangle = \langle B' \rangle + \langle B'' \rangle = W' + W'' = V$, also bildet $B' \cup B''$ ein Erzeugendensystem von V . Weiters haben wir $B' \cap B'' \subseteq W' \cap W'' = \{0\}$ und daher $B' \cap B'' = \emptyset$, denn jeder Basisvektor muss verschieden von Null sein. Um die lineare Unabhängigkeit von $B' \cup B''$ einzusehen, sei nun

$$\sum_{b \in B' \cup B''} \lambda_b b = 0,$$

wobei die Skalare λ_b fast alle verschwinden. Da $B' \cap B'' = \emptyset$, folgt

$$\underbrace{\sum_{b \in B'} \lambda_b b}_{\in W'} = - \underbrace{\sum_{b \in B''} \lambda_b b}_{\in W''},$$

also $\sum_{b \in B'} \lambda_b b = 0 = \sum_{b \in B''} \lambda_b b$, denn beide Seiten obiger Gleichung liegen in $W' \cap W'' = \{0\}$. Da B' und B'' beide linear unabhängig sind, folgt $\lambda_b = 0$, für alle $b \in B' \cup B''$. Somit ist $B' \cup B''$ linear unabhängig, also eine Basis von V . \square

Wir wollen nun zeigen, dass jeder Vektorraum eine Basis besitzt. Eine Methode Basen zu konstruieren besteht darin Erzeugendensysteme solange zu verkleinern, bis man bei einem linear unabhängigen Erzeugendensystem, also einer Basis, anlangt. Ist etwa E ein linear abhängiges Erzeugendensystem von V , dann existiert nach Proposition III.2.6 eine echte Teilmenge $E' \subsetneq E$, sodass $\langle E' \rangle = \langle E \rangle = V$. Wir erhalten also ein echt kleineres Erzeugendensystem E' von V . Ist E' immer noch linear abhängig können wir die selbe Konstruktion auf E' anwenden und erhalten ein noch kleineres Erzeugendensystem $E'' \subsetneq E'$ von V . Im Allgemeinen wird dieser Prozess nicht abbrechen, und liefert auch keine Basis von V . War das ursprüngliche Erzeugendensystem jedoch endlich, so müssen wir nach endlich vielen Schritten ein linear unabhängiges Erzeugendensystem, d.h. eine Basis von V erhalten. Wir fassen diese Überlegungen in folgender Proposition zusammen.

III.3.17. PROPOSITION. *Ist E ein endliches Erzeugendensystem eines Vektorraums V , dann existiert eine Basis B von V mit $B \subseteq E$. Insbesondere besitzt jeder endlich erzeugte Vektorraum eine Basis.*

Ein anderer Ansatz zur Konstruktion von Basen besteht darin linear unabhängige Teilmengen durch Hinzunahme von Vektoren so zu erweitern, dass die entstehenden Mengen immer noch linear unabhängig sind. Dazu:

III.3.18. LEMMA. *Sei A eine linear unabhängige Teilmenge eines Vektorraums V und $v \in V \setminus \langle A \rangle$. Dann ist auch $A \cup \{v\}$ linear unabhängig.*

BEWEIS. Seien also $a_1, \dots, a_n \in A$ paarweise verschieden und $\lambda, \lambda_1, \dots, \lambda_n \in \mathbb{K}$ so, dass

$$0 = \lambda v + \lambda_1 a_1 + \dots + \lambda_n a_n.$$

Wäre $\lambda \neq 0$, erhielten wir $v = (-\lambda^{-1}\lambda_1)a_1 + \dots + (-\lambda^{-1}\lambda_n)a_n \in \langle A \rangle$, ein Widerspruch zur Voraussetzung $v \in V \setminus \langle A \rangle$. Es muss daher $\lambda = 0$ gelten. Somit ist auch $0 = \lambda_1 a_1 + \dots + \lambda_n a_n$. Da A linear unabhängig ist, folgt $\lambda_1 = \dots = \lambda_n = 0$. Dies zeigt, dass $A \cup \{v\}$ linear unabhängig ist. \square

Ist nun A eine linear unabhängige Teilmenge eines Vektorraums V , die nicht ganz V aufspannt, dann existiert nach Lemma III.3.18 eine echt größere linear unabhängige Teilmenge $A' \supsetneq A$. Ist A' ein Erzeugendensystem von V , dann haben wir eine Basis gefunden. Ist A' kein Erzeugendensystem, dann erhalten wir mittels Lemma III.3.18 eine noch größere linear unabhängige Teilmenge $A'' \supsetneq A'$. Dieser Prozess lässt sich fortsetzen, er wird i.A. aber nicht abbrechen, und es ist an dieser Stelle nicht klar wie sich daraus Basen von V gewinnen lassen. Lemma III.3.18 erlaubt es jedoch Basen als maximal linear unabhängige Teilmengen zu charakterisieren.

III.3.19. PROPOSITION. *Für eine Teilmenge B eines Vektorraums V sind folgende Aussagen äquivalent:*

- (a) B ist eine Basis von V .
- (b) B ist ein minimales Erzeugendensystem von V , d.h. $\langle B \rangle = V$ und für jede echte Teilmenge $A \subsetneq B$ ist $\langle A \rangle \neq V$.
- (c) B ist eine maximal linear unabhängige Teilmenge von V , d.h. B ist linear unabhängig und jede echte Obermenge A , d.h. $B \subsetneq A \subseteq V$, ist linear abhängig.

BEWEIS. Die Äquivalenz (a) \Leftrightarrow (b) folgt sofort aus Proposition III.2.11.

Ad (a) \Rightarrow (c): Sei also B eine Basis von V . Insbesondere ist B linear unabhängig. Da B ein Erzeugendensystem von V bildet, gilt für jede Obermenge A von B , d.h. $B \subseteq A \subseteq V$, auch $\langle A \rangle = \langle B \rangle$, denn $V = \langle B \rangle \subseteq \langle A \rangle \subseteq V$. Nach Proposition III.2.6 ist daher jede echte Obermenge, $B \subsetneq A \subseteq V$, linear abhängig.

Ad (c) \Rightarrow (a): Sei also B eine maximal linear unabhängige Teilmenge von V . Jede echte Obermenge A mit $B \subsetneq A \subseteq V$ ist daher linear abhängig. Insbesondere ist für jedes $v \in V \setminus B$ die Menge $B \cup \{v\}$ linear abhängig, also $v \in \langle B \rangle$ nach Lemma III.3.18. Dies zeigt $V \setminus B \subseteq \langle B \rangle$. Da auch $B \subseteq \langle B \rangle$ gilt, erhalten wir $V = B \cup (V \setminus B) \subseteq \langle B \rangle \subseteq V$, also $\langle B \rangle = V$. Somit ist B ein Erzeugendensystem von V , also eine Basis. \square

Um nun die Existenz von Basen zeigen zu können, benötigen wir ein Hilfsmittel aus der Mengenlehre. Wir wiederholen zunächst die notwendigen Begriffe.

Sei X eine Menge und \leq eine Halbordnung auf X , d.h. eine reflexive, transitive und antisymmetrische Relation auf X , siehe [7, Kapitel 4]. Eine Teilmenge $A \subseteq X$ heißt nach *oben beschränkt* wenn $s \in X$ existiert, sodass $a \leq s$, für alle $a \in A$. Jedes solche s wird *eine obere Schranke* von A genannt. Unter einer *Kette in X* verstehen wir eine totalgeordnete Teilmenge $K \subseteq X$, d.h. für je zwei $k_1, k_2 \in K$ gilt stets $k_1 \leq k_2$ oder $k_2 \leq k_1$. Ein Element $x_0 \in X$ wird *maximal* genannt, wenn für jedes $x \in X$ mit $x_0 \leq x$ schon $x_0 = x$ gilt. In anderen Worten, ein Element von X ist genau dann maximal, wenn kein echt größeres Element von X existiert. Das wesentliche Hilfsmittel aus der Mengenlehre, das wir zur Konstruktion von Basen allgemeiner Vektorräume verwenden werden ist das folgende nach Max Zorn benannte Lemma.

III.3.20. LEMMA (Lemma von Zorn). *Jede halbgeordnete Menge, in der jede Kette eine obere Schranke hat, besitzt mindestens ein maximales Element.*

III.3.21. BEMERKUNG. Das Lemma von Zorn ist zum Auswahlaxiom äquivalent, siehe etwa [4, letztes Kapitel]. Dieses besagt, dass es stets möglich ist bei einer Familie nicht leerer Mengen $X_i, i \in I$, aus jeder der Mengen X_i ein Element auszuwählen, in anderen Worten, es existiert eine Abbildung $\alpha: I \rightarrow \bigcup_{i \in I} X_i$, sodass $\alpha(i) \in X_i$, für jedes $i \in I$. Dies bedeutet gerade, dass das Produkt $\prod_{i \in I} X_i$ nicht leer ist. Äquivalent dazu ist auch folgende Aussage: Jede surjektive Abbildung $f: X \rightarrow Y$ besitzt ein Rechtsinverses, d.h. es existiert eine Abbildung $g: Y \rightarrow X$, sodass $f \circ g = \text{id}_Y$. Es soll an dieser Stelle auch erwähnt sein, dass das Auswahlaxiom unabhängig vom Zermelo–Fraenkel Axiomensystem (ZFA) der Mengenlehre ist. Ist ZFA widerspruchsfrei, dann gilt dies auch für ZFA+Auswahlaxiom, aber auch für ZFA+(Negation des Auswahlaxioms). Die erste Aussage wurde von Kurt Gödel 1938 bewiesen, die zweite von Paul Cohen im Jahre 1963.

III.3.22. SATZ (Existenz von Basen). *Sei V ein Vektorraum, $A \subseteq V$ eine linear unabhängige Teilmenge und E ein Erzeugendensystem von V , sodass $A \subseteq E$. Dann existiert eine Basis B von V mit $A \subseteq B \subseteq E$.*

BEWEIS. Betrachte folgende Teilmenge der Potenzmenge von V ,

$$X := \{M \mid A \subseteq M \subseteq E \text{ und } M \text{ ist linear unabhängig}\}.$$

Die Mengeninklusion definiert eine Halbordnung \subseteq auf X . Wir wollen nun zeigen, dass sich das Lemma von Zorn auf X anwenden lässt. Sei also K eine Kette in X . Es genügt zu zeigen, dass

$$S := \bigcup_{M \in K} M$$

in X liegt, denn dann ist S offenbar eine obere Schranke für K . Offensichtlich gilt $A \subseteq S \subseteq E$, denn $A \subseteq M \subseteq E$ für jedes $M \in K$, also auch für deren Vereinigung.

Es bleibt daher nur noch die lineare Unabhängigkeit von S zu zeigen. Seien dazu $s_1, \dots, s_n \in S$ paarweise verschieden und $\lambda_1, \dots, \lambda_n \in \mathbb{K}$, sodass $\lambda_1 s_1 + \dots + \lambda_n s_n = 0$. Nach Konstruktion von S existiert zu jedem $i \in \{1, \dots, n\}$ ein $M_i \in K$, sodass $s_i \in M_i$. Da endliche totalgeordnete Mengen stets ein (eindeutiges) Maximum besitzen, vgl. Übungsaufgabe 63, existiert $i_0 \in \{1, \dots, n\}$, sodass $M_i \subseteq M_{i_0}$ für alle $i = 1, \dots, n$. Die Vektoren s_1, \dots, s_n liegen daher alle in M_{i_0} . Da M_{i_0} linear unabhängig ist, folgt $\lambda_1 = \dots = \lambda_n = 0$. Dies zeigt, dass S linear unabhängig ist.

Wir haben oben gesehen, dass jede Kette in X nach oben beschränkt ist. Nach dem Lemma von Zorn existiert daher ein maximales Element $B \in X$. Nach Konstruktion ist B eine linear unabhängige Teilmenge von V und es gilt $A \subseteq B \subseteq E$. Wegen der Maximalität von B ist jede Obermenge B' mit $B \subsetneq B' \subseteq E$ linear abhängig. Für jedes $e \in E \setminus B$ ist daher $B \cup \{e\}$ linear abhängig, also $e \in \langle B \rangle$ nach Lemma III.3.18. Dies zeigt $E \setminus B \subseteq \langle B \rangle$. Zusammen mit $B \subseteq \langle B \rangle$ folgt $E \subseteq \langle B \rangle$. Da E ein Erzeugendensystem von V bildet, erhalten wir $V = \langle E \rangle \subseteq \langle B \rangle \subseteq V$, also $\langle B \rangle = V$. Dies zeigt, dass B ein Erzeugendensystem von V bildet, also ist B die gesuchte Basis. \square

III.3.23. KOROLLAR.

- (a) Jeder Vektorraum besitzt eine Basis.
- (b) Jede linear unabhängige Teilmenge kann zu einer Basis erweitert werden.
- (c) Jedes Erzeugendensystem enthält eine Basis.

BEWEIS. Dies folgt aus Satz III.3.22 mit $A = \emptyset$ bzw. $E = V$. \square

III.3.24. BEMERKUNG. Es lässt sich zeigen, dass je zwei Basen eines Vektorraums gleiche Kardinalität haben, d.h. bijektiv aufeinander abgebildet werden können, siehe etwa [6, Theorem 1.12]. Den endlich-dimensionalen Fall werden wir in Kapitel IV genau diskutieren.

III.3.25. KOROLLAR. Ist W ein Teilraum eines Vektorraums V , dann lässt sich jede lineare Abbildung $\varphi: W \rightarrow U$ zu einer linearen Abbildung $\tilde{\varphi}: V \rightarrow U$ fortsetzen, d.h. $\tilde{\varphi}|_W = \varphi$.

BEWEIS. Nach Korollar III.3.23(a) besitzt W eine Basis B . Offensichtlich ist B auch in V linear unabhängig. Nach Korollar III.3.23(b) existiert daher eine Basis \tilde{B} von V mit $B \subseteq \tilde{B}$. Nach Proposition III.3.2 existiert eine lineare Abbildung $\tilde{\varphi}: V \rightarrow U$, sodass $\tilde{\varphi}|_B = \varphi|_B$ und $\tilde{\varphi}|_{\tilde{B} \setminus B} = 0$. Nach Proposition III.3.2 gilt $\tilde{\varphi}|_W = \varphi$, also ist $\tilde{\varphi}$ die gesuchte Fortsetzung von φ . \square

III.3.26. KOROLLAR. Jeder Teilraum besitzt ein Komplement.

BEWEIS. Sei also W Teilraum eines Vektorraums V . Nach Korollar III.3.25 existiert eine lineare Abbildung $\pi: V \rightarrow W$, sodass $\pi|_W = \text{id}_W$. Daraus folgt $\text{img}(\pi) = W$ und $\pi \circ \pi = \pi$. Nach Proposition II.5.8 ist daher $\ker(\pi)$ ein Komplement von W in V , d.h. $V = W \oplus \ker(\pi)$. \square

III.3.27. KOROLLAR.

- (a) Jede injektive lineare Abbildung $\varphi: V \rightarrow W$ besitzt eine lineare Linksinverse, d.h. es existiert eine lineare Abbildung $\psi: W \rightarrow V$ mit $\psi \circ \varphi = \text{id}_V$.
- (b) Jede surjektive lineare Abbildung $\varphi: V \rightarrow W$ besitzt eine lineare Rechtsinverse, d.h. es existiert eine lineare Abbildung $\psi: W \rightarrow V$ mit $\varphi \circ \psi = \text{id}_W$.

BEWEIS. Ad (a): Wegen der Injektivität von φ ist $\varphi: V \rightarrow \text{img}(\varphi)$ eine lineare Bijektion, also ein Isomorphismus. Nach Korollar III.3.25 lässt sich die lineare Abbildung $\varphi^{-1}: \text{img}(\varphi) \rightarrow V$ zu einer linearen Abbildung $\psi: W \rightarrow V$ fortsetzen, d.h. $\psi|_{\text{img}(\varphi)} = \varphi^{-1}$. Es folgt nun $\psi \circ \varphi = \psi|_{\text{img}(\varphi)} \circ \varphi = \varphi^{-1} \circ \varphi = \text{id}_V$.

Ad (b): Sei B eine Basis von W . Wegen der Surjektivität von φ existiert eine Abbildung $f: B \rightarrow V$, sodass $\varphi \circ f = \text{id}_B$. Nach Proposition III.3.2 existiert eine lineare Abbildung $\psi: W \rightarrow V$, sodass $\psi|_B = f$. Es folgt $\varphi \circ \psi|_B = \varphi \circ f = \text{id}_B$ und nach der Eindeutigkeitsaussage in Proposition III.3.2(c) daher $\varphi \circ \psi = \text{id}_W$. \square

III.3.28. KOROLLAR. Sei V ein \mathbb{K} -Vektorraum und $0 \neq v \in V$. Dann existiert ein lineares Funktional $\alpha: V \rightarrow \mathbb{K}$ mit $\alpha(v) = 1$.

BEWEIS. Da $v \neq 0$, ist die lineare Abbildung $\varphi: \mathbb{K} \rightarrow V$, $\varphi(\lambda) := \lambda v$, injektiv. Nach Korollar III.3.27(a) existiert daher eine lineare Abbildung $\alpha: V \rightarrow \mathbb{K}$, sodass $\alpha \circ \varphi = \text{id}_{\mathbb{K}}$. Es gilt daher $\alpha(v) = \alpha(\varphi(1)) = (\alpha \circ \varphi)(1) = \text{id}_{\mathbb{K}}(1) = 1$. \square

III.3.29. PROPOSITION. Seien W' und W'' zwei Teilräume eines Vektorraums V . Dann existieren eine Basis B' von W' und eine Basis B'' von W'' , sodass $B' \cap B''$ eine Basis von $W' \cap W''$ bildet, und $B' \cup B''$ eine Basis von $W' + W''$ bildet.

BEWEIS. Nach Korollar III.3.26 existiert ein zu $W := W' \cap W''$ komplementärer Teilraum U in $W' + W''$, d.h.

$$W' + W'' = W \oplus U.$$

Für die Teilräume $U' := W' \cap U$ und $U'' := W'' \cap U$ gilt dann:

$$W' = W \oplus U', \quad W'' = W \oplus U'', \quad U = U' \oplus U''.$$

Es ist nämlich $W \cap U' \subseteq W \cap U = \{0\}$ und $W + U' \subseteq W'$ aber auch $W' \subseteq W + U'$, denn zu jedem $w' \in W'$ existieren $w \in W$ und $u \in U$ mit $w' = w + u$, da aber $u = w' - w \in W'$ folgt $u \in U'$, also $w' \in W + U'$. Dies zeigt die erste Behauptung, $W' = W \oplus U'$. Die zweite, $W'' = W \oplus U''$, lässt sich analog beweisen. Daraus folgt weiters $W' + W'' = W + U' + U''$ und somit $U \subseteq U' + U''$, denn zu jedem $u \in U$ existieren $w \in W$, $u' \in U'$ und $u'' \in U''$ mit $u = w + u' + u''$, da aber $w = u - u' - u'' \in W \cap U = \{0\}$ muss $w = 0$ gelten, also $u = u' + u'' \in U' + U''$. Da offensichtlich $U' + U'' \subseteq U$ erhalten wir $U = U' + U''$. Schließlich haben wir auch $U' \cap U'' \subseteq W' \cap W'' \cap U = W \cap U = \{0\}$, also $U = U' \oplus U''$.

Nach Korollar III.3.23(a) existiert eine Basis B von W , eine Basis C' von U' und eine Basis C'' von U'' . Nach Proposition III.3.16 ist $B' := B \cup C'$ eine Basis von W' , $B'' := B \cup C''$ eine Basis von W'' , $C' \cup C''$ eine Basis von U , $B' \cup B'' = B \cup C' \cup C''$ eine Basis von $W' + W''$ und $C' \cap C'' = \emptyset$, also auch $B' \cap B'' = B$ eine Basis von $W' \cap W'' = W$. \square

III.4. Dualräume. Eine Möglichkeit einen Teilraum W eines \mathbb{K} -Vektorraums V zu beschreiben besteht darin eine Basis, oder allgemeiner, ein Erzeugendensystem von W anzugeben, siehe oben. Teilräume lassen sich aber auch durch lineare Gleichungen beschreiben. Eine naheliegende Verallgemeinerung einer Gleichung $a_1x_1 + \cdots + a_nx_n = 0$ für Vektoren $x \in \mathbb{K}^n$ ist die Bedingung $\alpha(v) = 0$, wobei $\alpha: V \rightarrow \mathbb{K}$ linear ist. Haben wir eine Menge linearer Abbildungen $\alpha_i: V \rightarrow \mathbb{K}$, $i \in I$, dann bildet

$$\{v \in V \mid \forall i \in I : \alpha_i(v) = 0\} = \bigcap_{i \in I} \ker(\alpha_i)$$

einen Teilraum von V , den wir als Lösungsraum des Gleichungssystems $\alpha_i(v) = 0$, $i \in I$, verstehen können. Wir werden unten sehen, dass sich jeder Teilraum von V in dieser Form schreiben lässt. Für eine effektive Beschreibung sind natürlich möglichst kleine Gleichungssysteme interessant. Für endlich-dimensionale Vektorräume werden wir diesen Aspekt im nächsten Kapitel genau behandeln.

III.4.1. DEFINITION (Dualraum). Sei V ein Vektorraum über \mathbb{K} . Unter einem *linearen Funktional* auf V verstehen wir eine lineare Abbildung $V \rightarrow \mathbb{K}$. Unter dem *Dualraum* von V verstehen wir den Vektorraum $V^* := L(V, \mathbb{K})$, d.h. den Vektorraum aller linearen Funktionale auf V , vgl. Proposition II.3.18.

III.4.2. BEISPIEL. Es ist $(\mathbb{K}^n)^* = L(\mathbb{K}^n, \mathbb{K}) \cong M_{1 \times n}(\mathbb{K})$, vgl. Satz II.4.4, lineare Funktionale auf \mathbb{K}^n können daher durch Zeilenvektoren beschrieben werden.

III.4.3. PROPOSITION. Jede Abbildung $\varphi: V \rightarrow W$ induziert eine lineare Abbildung zwischen den Dualräumen, $\varphi^t: W^* \rightarrow V^*$, $\varphi^t(\beta) := \beta \circ \varphi$, $\beta \in W^*$. Die Zuordnung

$$L(V, W) \rightarrow L(W^*, V^*), \quad \varphi \mapsto \varphi^t, \quad (\text{III.1})$$

ist linear und injektiv, es gilt daher

$$(\varphi_1 + \varphi_2)^t = \varphi_1^t + \varphi_2^t \quad \text{sowie} \quad (\lambda\varphi)^t = \lambda\varphi^t,$$

für beliebige lineare Abbildungen $\varphi, \varphi_1, \varphi_2: V \rightarrow W$ und $\lambda \in \mathbb{K}$. Für jede weitere lineare Abbildung $\psi: W \rightarrow U$ gilt

$$(\psi \circ \varphi)^t = \varphi^t \circ \psi^t \quad \text{und} \quad \text{id}_V^t = \text{id}_{V^*}.$$

Ist $\varphi: V \rightarrow W$ ein Isomorphismus, so ist auch $\varphi^t: W^* \rightarrow V^*$ ein Isomorphismus,

$$(\varphi^t)^{-1} = (\varphi^{-1})^t.$$

Insbesondere folgt aus $V \cong W$ auch $V^* \cong W^*$.

BEWEIS. Beachte zunächst, dass $\varphi^t(\beta) = \beta \circ \varphi$ als Komposition linearer Abbildungen selbst linear ist, d.h. $\varphi^t(\beta) \in V^*$ für jedes $\beta \in W^*$. Die Linearität von $\varphi^t: W^* \rightarrow V^*$ folgt aus Proposition II.3.18(a), denn $\varphi^t(\beta_1 + \beta_2) = (\beta_1 + \beta_2) \circ \varphi = \beta_1 \circ \varphi + \beta_2 \circ \varphi = \varphi^t(\beta_1) + \varphi^t(\beta_2)$ und $\varphi^t(\lambda\beta) = (\lambda\beta) \circ \varphi = \lambda(\beta \circ \varphi) = \lambda\varphi^t(\beta)$, für beliebige $\beta, \beta_1, \beta_2 \in W^*$ und $\lambda \in \mathbb{K}$. Aus Proposition II.3.18(b) erhalten wir weiters $(\varphi_1 + \varphi_2)^t(\beta) = \beta \circ (\varphi_1 + \varphi_2) = \beta \circ \varphi_1 + \beta \circ \varphi_2 = \varphi_1^t(\beta) + \varphi_2^t(\beta) = (\varphi_1^t + \varphi_2^t)(\beta)$

und $(\lambda\varphi)^t(\beta) = \beta \circ (\lambda\varphi) = \lambda(\beta \circ \varphi) = \lambda\varphi^t(\beta)$, für jedes $\beta \in W^*$ und $\lambda \in \mathbb{K}$. Die Zuordnung (III.1) ist daher linear. Ist $0 \neq \varphi \in L(V, W)$, dann existiert $v \in V$ mit $\varphi(v) \neq 0$. Nach Korollar III.3.28 existiert daher ein Funktional $\beta \in W^*$ mit $\beta(\varphi(v)) = 1$. Es folgt $\varphi^t(\beta)(v) = \beta(\varphi(v)) = 1 \neq 0$, also $\varphi^t(\beta) \neq 0$ und daher $\varphi^t \neq 0$. Dies zeigt, dass die Abbildung (III.1) trivialen Kern hat, und daher injektiv ist, vgl. Proposition II.3.22. Weiters gilt $\text{id}_V^t(\alpha) = \alpha \circ \text{id}_V = \alpha$, für jedes $\alpha \in V^*$, also $\text{id}_V^t = \text{id}_{V^*}$, sowie $(\psi \circ \varphi)^t(\gamma) = \gamma \circ (\psi \circ \varphi) = (\gamma \circ \psi) \circ \varphi = \varphi^t(\gamma \circ \psi) = \varphi^t(\psi^t(\gamma)) = (\varphi^t \circ \psi^t)(\gamma)$, für jedes $\gamma \in U^*$, also $(\psi \circ \varphi)^t = \varphi^t \circ \psi^t$. Ist $\varphi: V \rightarrow W$ ein Isomorphismus, so folgt $\varphi^t \circ (\varphi^{-1})^t = (\varphi^{-1} \circ \varphi)^t = \text{id}_V^t = \text{id}_{V^*}$ und $(\varphi^{-1})^t \circ \varphi^t = (\varphi \circ \varphi^{-1})^t = \text{id}_W^t = \text{id}_{W^*}$, d.h. φ^t ist invertierbar mit Umkehrabbildung $(\varphi^t)^{-1} = (\varphi^{-1})^t$. \square

III.4.4. BEMERKUNG. Die Abbildung (III.1) ist i.A. nicht surjektiv.

III.4.5. BEMERKUNG. Nach Satz II.4.4 ist $\phi: \mathbb{K}^n \xrightarrow{\cong} L(\mathbb{K}^n, \mathbb{K}) = (\mathbb{K}^n)^*$, $\phi(x) := \psi_{x^t}$, ein Isomorphismus. Ist $A \in M_{m \times n}(\mathbb{K})$ und $\psi_A: \mathbb{K}^n \rightarrow \mathbb{K}^m$, $\psi_A(x) = Ax$, die damit assoziierte lineare Abbildung, dann kommutiert das folgende Diagramm

$$\begin{array}{ccc} \mathbb{K}^m & \xrightarrow[\cong]{\phi_{\mathbb{K}^m}} & (\mathbb{K}^m)^* \\ \psi_{A^t} \downarrow & & \downarrow (\psi_A)^t \\ \mathbb{K}^n & \xrightarrow[\cong]{\phi_{\mathbb{K}^n}} & (\mathbb{K}^n)^* \end{array} \quad \text{d.h.} \quad \phi_{\mathbb{K}^n} \circ \psi_{A^t} = (\psi_A)^t \circ \phi_{\mathbb{K}^m}.$$

Es gilt nämlich $((\psi_A)^t \circ \phi_{\mathbb{K}^m})(x) = (\psi_A)^t(\psi_{x^t}) = \psi_{x^t} \circ \psi_A = \psi_{x^t A} = \psi_{(A^t x)^t} = \phi_{\mathbb{K}^n}(A^t x) = (\phi_{\mathbb{K}^n} \circ \psi_{A^t})(x)$, für alle $x \in \mathbb{K}^m$. Bis auf den Isomorphismus ϕ ist die duale Abbildung, $(\psi_A)^t$, daher durch die transponierte Matrix, A^t , gegeben, vgl. Proposition II.4.11.

III.4.6. DEFINITION (Annihilator). Unter dem *Annihilator* einer Teilmenge A eines \mathbb{K} -Vektorraums V verstehen wir den Teilraum

$$A^\circ := \{\alpha \in V^* : \alpha|_A = 0\} \subseteq V^*.$$

III.4.7. LEMMA. Sind A und B Teilmengen eines Vektorraums V , dann gilt:

- (a) $\{0\}^\circ = V^*$ und $V^\circ = \{0\}$.
- (b) $A \subseteq B \Rightarrow B^\circ \subseteq A^\circ$.
- (c) $(A \cup B)^\circ = A^\circ \cap B^\circ$.
- (d) $A^\circ = \langle A \rangle^\circ$.
- (e) $\langle A \rangle = \bigcap_{\alpha \in A^\circ} \ker(\alpha)$.
- (f) $\langle A \rangle \subseteq \langle B \rangle \Leftrightarrow B^\circ \subseteq A^\circ$.
- (g) $\langle A \rangle = \langle B \rangle \Leftrightarrow A^\circ = B^\circ$.
- (h) Für jeden Teilraum W von V gilt: $W = \bigcap_{\alpha \in W^\circ} \ker(\alpha)$.
- (i) Für je zwei Teilräume W_1 und W_2 von V gilt: $W_1 \subseteq W_2 \Leftrightarrow W_2^\circ \subseteq W_1^\circ$.
- (j) Für je zwei Teilräume W_1 und W_2 von V gilt: $W_1 = W_2 \Leftrightarrow W_1^\circ = W_2^\circ$.

BEWEIS. Die Behauptungen (a), (b) und (c) sind trivial. Behauptung (d) folgt aus Lemma III.1.4(g). Ad (e): Offensichtlich gilt $A \subseteq \bigcap_{\alpha \in A^\circ} \ker(\alpha)$ und daher auch $\langle A \rangle \subseteq \bigcap_{\alpha \in A^\circ} \ker(\alpha)$, denn als Durchschnitt von Teilräumen bildet die rechte Seite einen Teilraum von V . Für die umgekehrte Inklusion sei nun $v \in V \setminus \langle A \rangle$. Es genügt ein lineares Funktional $\alpha: V \rightarrow \mathbb{K}$ zu konstruieren, sodass $\alpha(v) = 1$ und $\alpha|_A = 0$. Für die Konstruktion eines solchen Funktionals α bezeichne $\pi: V \rightarrow V/\langle A \rangle$ die kanonische Projektion. Da $v \notin \langle A \rangle$ gilt $\pi(v) \neq 0$. Nach Korollar III.3.28 existiert ein lineares Funktional $\bar{\alpha}: V/\langle A \rangle \rightarrow \mathbb{K}$ mit $\bar{\alpha}(\pi(v)) = 1$. Für das Funktional $\alpha := \bar{\alpha} \circ \pi \in V^*$ gilt daher $\alpha(v) = 1$ aber auch $\alpha|_A = 0$, denn $A \subseteq \ker(\pi)$. Ad (f): Die Implikation $\langle A \rangle \subseteq \langle B \rangle \Rightarrow B^\circ \subseteq A^\circ$ folgt aus (b) und (d), die umgekehrte Implikation aus (e). Aus (f) erhalten wir auch sofort (g). Die verbleibenden Behauptungen (h), (i) und (j) folgen aus (e), (f) und (g), denn für jeden Teilraum W gilt $W = \langle W \rangle$. \square

III.4.8. BEMERKUNG. Ist W ein Teilraum eines Vektorraums V und $A \subseteq W^\circ$ ein Erzeugendensystem, d.h. $\langle A \rangle = W^\circ$, dann gilt

$$W = \{v \in V \mid \forall \alpha \in A : \alpha(v) = 0\} = \bigcap_{\alpha \in A} \ker(\alpha),$$

jedes Erzeugendensysteme von W° liefert daher ein Gleichungssysteme für W . Dies folgt aus Lemma III.4.7(h) und

$$\bigcap_{\alpha \in A} \ker(\alpha) = \bigcap_{\alpha \in \langle A \rangle} \ker(\alpha).$$

Um die letzte Gleichung einzusehen, sei $v \in \bigcap_{\alpha \in A} \ker(\alpha)$. Für beliebige $\alpha_i \in A$ und $\lambda_i \in \mathbb{K}$ gilt dann auch $(\lambda_1 \alpha_1 + \dots + \lambda_n \alpha_n)(v) = \lambda_1 \alpha_1(v) + \dots + \lambda_n \alpha_n(v) = 0$. Dies zeigt $\bigcap_{\alpha \in A} \ker(\alpha) \subseteq \bigcap_{\alpha \in \langle A \rangle} \ker(\alpha)$, die andere Inklusion ist trivial. Umgekehrt folgt i.A. aus $W = \bigcap_{\alpha \in A} \ker(\alpha)$ nicht, dass A ein Erzeugendensystem von W bildet. Betrachten wir etwa den Vektorraum $V = \mathbb{K}[z]$ und bezeichnet $\alpha_i \in \mathbb{K}[z]^*$ das lineare Funktional, das einem Polynom seinen i -ten Koeffizienten zuordnet, d.h. $\alpha_i(p_0 + p_1 z + p_2 z^2 + \dots) := p_i$, dann gilt offensichtlich $\bigcap_i \ker(\alpha_i) = \{0\}$, aber die Menge der Funktionale $\{\alpha_0, \alpha_1, \dots\}$ bildet kein Erzeugendensystem von $\{0\}^\circ = \mathbb{K}[z]^*$, etwa lässt sich das lineare Funktional $\alpha \in \mathbb{K}[z]^*$, $\alpha(p_0 + p_1 z + p_2 z^2 + \dots) := \sum_i p_i$, nicht als Linearkombination der Funktionale α_i schreiben.

III.4.9. SATZ. Für je zwei Teilräume W_1 und W_2 eines Vektorraums V gilt

$$(W_1 + W_2)^\circ = W_1^\circ \cap W_2^\circ \quad \text{und} \quad (W_1 \cap W_2)^\circ = W_1^\circ + W_2^\circ.$$

Insbesondere haben wir:

$$\begin{aligned} W_1^\circ \cap W_2^\circ = \{0\} &\Leftrightarrow W_1 + W_2 = V \\ W_1^\circ + W_2^\circ = V^* &\Leftrightarrow W_1 \cap W_2 = \{0\} \\ W_1^\circ \oplus W_2^\circ = V^* &\Leftrightarrow W_1 \oplus W_2 = V \end{aligned}$$

BEWEIS. Die erste Gleichheit folgt aus:

$$\begin{aligned} W_1^\circ \cap W_2^\circ &= (W_1 \cup W_2)^\circ && \text{nach Lemma III.4.7(c)} \\ &= \langle W_1 \cup W_2 \rangle^\circ && \text{nach Lemma III.4.7(d)} \\ &= (W_1 + W_2)^\circ && \text{nach Lemma III.1.4(e)} \end{aligned}$$

Auch die Inklusion $W_1^\circ + W_2^\circ \subseteq (W_1 \cap W_2)^\circ$ lässt sich leicht zeigen: Aus $W_1 \cap W_2 \subseteq W_1$ folgt nämlich $W_1^\circ \subseteq (W_1 \cap W_2)^\circ$ und analog gilt $W_2^\circ \subseteq (W_1 \cap W_2)^\circ$. Somit ist $W_1^\circ \cup W_2^\circ \subseteq (W_1 \cap W_2)^\circ$. Da $W_1^\circ + W_2^\circ$ der kleinste Teilraum ist, der $(W_1 \cap W_2)^\circ$ enthält erhalten wir $W_1^\circ + W_2^\circ \subseteq (W_1 \cap W_2)^\circ$.

Um auch die umgekehrte Inklusion $(W_1 \cap W_2)^\circ \subseteq W_1^\circ + W_2^\circ$ zu zeigen sei nun $\alpha \in (W_1 \cap W_2)^\circ$, d.h. $\alpha \in V^*$ und $W_1 \cap W_2 \subseteq \ker(\alpha)$. Wir werden unten eine lineare Abbildung ρ mit folgenden Eigenschaften konstruieren:

$$\rho: V \rightarrow V, \quad \rho|_{W_2} = \text{id}_{W_2}, \quad \rho(W_1) \subseteq W_1 \cap W_2.$$

Es ist dann $\alpha_1 := \alpha \circ \rho \in W_1^\circ$, denn für jedes $w_1 \in W_1$ gilt $\alpha_1(w_1) = \alpha(\rho(w_1)) = 0$, da ja $\rho(w_1) \in \rho(W_1) \subseteq W_1 \cap W_2 \subseteq \ker(\alpha)$. Setzen wir $\alpha_2 := \alpha - \alpha_1$ dann gilt auch $\alpha_2 \in W_2^\circ$, denn für jedes $w_2 \in W_2$ ist $\alpha_2(w_2) = \alpha(w_2) - \alpha_1(w_2) = \alpha(w_2) - \alpha(\rho(w_2)) = \alpha(w_2) - \alpha(w_2) = 0$. Somit erhalten wir $\alpha = \alpha_1 + \alpha_2 \in W_1^\circ + W_2^\circ$.

Für die Konstruktion von ρ wählen wir Basen B_1 von W_1 und B_2 von W_2 wie in Proposition III.3.29, d.h. $B_1 \cup B_2$ ist Basis von $W_1 + W_2$. Nach Proposition III.3.2 existiert eine lineare Abbildung $\bar{\rho}: W_1 + W_2 \rightarrow V$, sodass $\bar{\rho}|_{B_2} = \text{id}_{B_2}$ und $\bar{\rho}|_{B_1 \setminus B_2} = 0$. Wegen $B_1 = (B_1 \setminus B_2) \cup (B_1 \cap B_2)$ gilt daher auch $\bar{\rho}(B_1) \subseteq W_1 \cap W_2$. Aus der Eindeutigkeitsaussage in Proposition III.3.2(c) folgt daher $\bar{\rho}|_{W_2} = \text{id}_{W_2}$ und $\bar{\rho}(W_1) \subseteq W_1 \cap W_2$. Nach Korollar III.3.25 lässt sich $\bar{\rho}$ zu einer linearen Abbildung $\rho: V \rightarrow V$ erweitern, d.h. $\rho|_{W_1 + W_2} = \bar{\rho}$, und diese Fortsetzung hat die gewünschten Eigenschaften.

Die verbleibenden Behauptungen folgen nun aus Lemma III.4.7(a)&(j), denn $W_1 + W_2 = V \Leftrightarrow (W_1 + W_2)^\circ = V^\circ \Leftrightarrow W_1^\circ \cap W_2^\circ = \{0\}$ und $W_1 \cap W_2 = \{0\} \Leftrightarrow (W_1 \cap W_2)^\circ = \{0\}^\circ \Leftrightarrow W_1^\circ + W_2^\circ = V$. \square

III.4.10. SATZ. Für jede lineare Abbildung $\varphi: V \rightarrow W$ gilt

$$\ker(\varphi^t) = \text{img}(\varphi)^\circ \quad \text{und} \quad \text{img}(\varphi^t) = \ker(\varphi)^\circ.$$

Insbesondere ist φ genau dann injektiv, wenn φ^t surjektiv ist. Auch ist φ genau dann surjektiv, wenn φ^t injektiv ist.

BEWEIS. Es ist $\ker(\varphi^t) = \text{img}(\varphi)^\circ$, denn für $\beta \in W^*$ gilt:

$$\begin{aligned} \beta \in \ker(\varphi^t) &\Leftrightarrow \beta \circ \varphi = 0 \\ &\Leftrightarrow \forall v \in V : \beta(\varphi(v)) = 0 \\ &\Leftrightarrow \forall w \in \text{img}(\varphi) : \beta(w) = 0 \\ &\Leftrightarrow \beta|_{\text{img}(\varphi)} = 0 \\ &\Leftrightarrow \beta \in \text{img}(\varphi)^\circ \end{aligned}$$

Auch die Inklusion $\text{img}(\varphi^t) \subseteq \ker(\varphi)^\circ$ lässt sich leicht beweisen. Ist nämlich $\alpha \in \text{img}(\varphi^t) \subseteq V^*$ dann existiert $\beta \in W^*$ mit $\alpha = \varphi^t(\beta) = \beta \circ \varphi$, also gilt $\alpha(v) = (\beta \circ \varphi)(v) = \beta(\varphi(v)) = 0$ für jedes $v \in \ker(\varphi)$, und daher $\alpha \in \ker(\varphi)^\circ$.

Wir zeigen nun die umgekehrte Inklusion, $\ker(\varphi)^\circ \subseteq \text{img}(\varphi^t)$. Sei dazu $\alpha \in \ker(\varphi)^\circ$, d.h. $\alpha \in V^*$ und $\alpha|_{\ker(\varphi)} = 0$. Nach Satz II.6.3 existiert ein lineares Funktional $\bar{\alpha}: V/\ker(\varphi) \rightarrow \mathbb{K}$, sodass $\bar{\alpha} \circ \pi = \alpha$, wobei $\pi: V \rightarrow V/\ker(\varphi)$ die kanonische Projektion bezeichnet. Nach Korollar II.6.6 induziert φ einen linearen Isomorphismus $\bar{\varphi}: V/\ker(\varphi) \xrightarrow{\cong} \text{img}(\varphi)$ und es gilt $\bar{\varphi} \circ \pi = \varphi$. Nach Korollar III.3.25 lässt sich das lineare Funktional $\bar{\alpha} \circ \bar{\varphi}^{-1}: \text{img}(\varphi) \rightarrow \mathbb{K}$ zu einem linearen Funktional $\beta \in W^*$ fortsetzen, d.h. $\beta|_{\text{img}(\varphi)} = \bar{\alpha} \circ \bar{\varphi}^{-1}$. Nach Konstruktion gilt $\varphi^t(\beta) = \beta \circ \varphi = \beta|_{\text{img}(\varphi)} \circ \varphi = (\bar{\alpha} \circ \bar{\varphi}^{-1}) \circ (\bar{\varphi} \circ \pi) = \bar{\alpha} \circ \pi = \alpha$, also $\alpha \in \text{img}(\varphi^t)$. Damit ist auch $\text{img}(\varphi^t) = \ker(\varphi)^\circ$ gezeigt. Insbesondere folgt

$$\begin{aligned} \varphi \text{ ist surjektiv} &\Leftrightarrow \text{img}(\varphi) = W \\ &\Leftrightarrow \text{img}(\varphi)^\circ = W^\circ && \text{nach Lemma III.4.7(j)} \\ &\Leftrightarrow \text{img}(\varphi)^\circ = \{0\} && \text{nach Lemma III.4.7(a)} \\ &\Leftrightarrow \ker(\varphi^t) = \{0\} && \text{da } \ker(\varphi^t) = \text{img}(\varphi)^\circ \\ &\Leftrightarrow \varphi^t \text{ ist injektiv} && \text{nach Proposition II.3.22} \end{aligned}$$

und analog

$$\begin{aligned} \varphi \text{ ist injektiv} &\Leftrightarrow \ker(\varphi) = \{0\} && \text{nach Proposition II.3.22} \\ &\Leftrightarrow \ker(\varphi)^\circ = \{0\}^\circ && \text{nach Lemma III.4.7(j)} \\ &\Leftrightarrow \ker(\varphi)^\circ = V^* && \text{nach Lemma III.4.7(a)} \\ &\Leftrightarrow \text{img}(\varphi^t) = V^* && \text{da } \text{img}(\varphi^t) = \ker(\varphi)^\circ \\ &\Leftrightarrow \varphi^t \text{ ist surjektiv} \end{aligned}$$

Damit ist der Beweis des Satzes vollständig. \square

III.4.11. BEISPIEL. Sei $A \in M_{m \times n}(\mathbb{K})$, $\psi_A: \mathbb{K}^n \rightarrow \mathbb{K}^m$, $\psi_A(x) = Ax$, die damit assoziierte lineare Abbildung und $W = \ker(\psi_A)$ der Lösungsraum des homogenen Gleichungssystems $Ax = 0$. Bezeichnet $\alpha_i \in M_{1 \times n}(\mathbb{K}) = (\mathbb{K}^n)^*$ den i -ten Zeilenvektor von A , dann bildet $\{\alpha_1, \dots, \alpha_m\}$ ein Erzeugendensystem von W° . Da die Spalten von A^t ein Erzeugendensystem von $\text{img}(\psi_{A^t})$ bilden, folgt nämlich mit Bemerkung III.4.5, dass die Zeilen von A ein Erzeugendensystem von $\text{img}((\psi_A)^t)$ darstellen, nach Satz III.4.10 gilt jedoch $W^\circ = \ker(\psi_A)^\circ = \text{img}((\psi_A)^t)$.

III.4.12. KOROLLAR. *Ist W ein Teilraum eines Vektorraums V , dann gilt:*

(a) *Die kanonische Inklusion $\iota: W \rightarrow V$ induziert einen Isomorphismus*

$$V^*/W^\circ \cong W^*, \quad [\alpha] \mapsto \alpha|_W = \alpha \circ \iota, \quad \alpha \in V^*.$$

Wir können V^/W° daher in natürlicher Weise mit W^* identifizieren.*

(b) Die kanonische Projektion $\pi: V \rightarrow V/W$ induziert einen Isomorphismus

$$(V/W)^* \cong W^\circ, \quad \beta \mapsto \beta \circ \pi, \quad \beta \in (V/W)^*.$$

Wir können daher W° in natürlicher Weise mit $(V/W)^*$ identifizieren.

BEWEIS. Ad (a): Die Inklusionsabbildung $\iota: W \rightarrow V$ ist offensichtlich injektiv, und $\text{img}(\iota) = W$. Nach Satz III.4.10 ist daher die duale Abbildung $\iota^t: V^* \rightarrow W^*$ surjektiv, und $\ker(\iota^t) = \text{img}(\iota)^\circ = W^\circ$. Nach Korollar II.6.6 induziert ι^t also einen Isomorphismus $V^*/W^\circ \cong W^*$, $[\alpha] \mapsto \iota^t(\alpha) = \alpha \circ \iota = \alpha|_W$, wobei $\alpha \in V^*$.

Ad (b): Nach Satz II.6.3 ist die kanonische Projektion $\pi: V \rightarrow V/W$ surjektiv, und $\ker(\pi) = W$. Nach Satz III.4.10 ist daher die duale Abbildung $\pi^t: (V/W)^* \rightarrow V^*$ injektiv, und $\text{img}(\pi^t) = \ker(\pi)^\circ = W^\circ$. Somit liefert π^t einen Isomorphismus $(V/W)^* \cong W^\circ$, $\beta \mapsto \pi^t(\beta) = \beta \circ \pi$, wobei $\beta \in (V/W)^*$. \square

Ist V ein \mathbb{K} -Vektorraum und $v \in V$, dann ist die Abbildung

$$\text{ev}_v: V^* \rightarrow \mathbb{K}, \quad \text{ev}_v(\alpha) := \alpha(v),$$

linear, denn $\text{ev}_v(\alpha_1 + \alpha_2) = (\alpha_1 + \alpha_2)(v) = \alpha_1(v) + \alpha_2(v) = \text{ev}_v(\alpha_1) + \text{ev}_v(\alpha_2)$ und $\text{ev}_v(\lambda\alpha) = (\lambda\alpha)(v) = \lambda\alpha(v) = \lambda \text{ev}_v(\alpha)$, für beliebige $\alpha, \alpha_1, \alpha_2 \in V^*$ und $\lambda \in \mathbb{K}$, vgl. Beispiel II.3.8. Somit ist $\text{ev}_v \in (V^*)^*$. Der Vektorraum $(V^*)^*$ wird *Bidual* von V genannt und oft mit V^{**} bezeichnet. Die Zuordnung

$$\iota_V: V \rightarrow V^{**}, \quad \iota_V(v) := \text{ev}_v,$$

ist ebenfalls linear, denn es gilt $\text{ev}_{v_1+v_2}(\alpha) = \alpha(v_1+v_2) = \alpha(v_1) + \alpha(v_2) = \text{ev}_{v_1}(\alpha) + \text{ev}_{v_2}(\alpha) = (\text{ev}_{v_1} + \text{ev}_{v_2})(\alpha)$ und $\text{ev}_{\lambda v}(\alpha) = \alpha(\lambda v) = \lambda\alpha(v) = \lambda \text{ev}_v(\alpha) = (\lambda \text{ev}_v)(\alpha)$ für alle $v, v_1, v_2 \in V$ und $\lambda \in \mathbb{K}$.

III.4.13. PROPOSITION. Die oben besprochene kanonische lineare Abbildung

$$\iota_V: V \rightarrow V^{**}, \quad \iota_V(v)(\alpha) := \alpha(v), \quad v \in V, \alpha \in V^*,$$

ist injektiv. Für jeden Teilraum W von V gilt $\iota(W) \subseteq W^{\circ\circ}$. Für jede lineare Abbildung $\varphi: V \rightarrow U$ gilt $\varphi^{tt} \circ \iota_V = \iota_U \circ \varphi$, wobei $\varphi^{tt}: V^{**} \rightarrow U^{**}$ die zu φ^t duale Abbildung bezeichnet, d.h. $\varphi^{tt}(\xi)(\beta) = \xi(\varphi^t(\beta))$, $\xi \in V^{**}$, $\beta \in U^*$.

BEWEIS. Wir zeigen zunächst, dass ι_V injektiv ist. Sei dazu $0 \neq v \in V$. Nach Korollar III.3.28 existiert $\alpha \in V^*$ mit $\alpha(v) = 1$. Es gilt daher $\iota_V(v)(\alpha) = \alpha(v) = 1 \neq 0$, also $\iota_V(v) \neq 0$. Es folgt $\ker(\iota_V) = \{0\}$, also ist ι_V injektiv, siehe Proposition II.3.22.

Sei nun $W \subseteq V$ ein Teilraum, $w \in W$ und $\alpha \in W^\circ$, d.h. $\alpha \in V^*$ und $\alpha|_W = 0$. Dann folgt $\iota_V(w)(\alpha) = \alpha(w) = 0$. Da dies für alle $\alpha \in W^\circ$ gilt, folgt $\iota_V(w) \in W^{\circ\circ}$. Da dies für alle $w \in W$ richtig ist, erhalten wir $\iota_V(W) \subseteq W^{\circ\circ}$.

Sei schließlich $\varphi: V \rightarrow U$ linear, $v \in V$ und $\beta \in U^*$. Dann gilt

$$\begin{aligned}
 (\varphi^{tt} \circ \iota_V)(v)(\beta) &= \varphi^{tt}(\iota_V(v))(\beta) && \text{Definition der Komposition von Abb.} \\
 &= \iota_V(v)(\varphi^t(\beta)) && \text{Definition von der dualen Abb. } \varphi^{tt} \\
 &= \varphi^t(\beta)(v) && \text{Definition von } \iota_V \\
 &= \beta(\varphi(v)) && \text{Definition von der dualen Abb. } \varphi^t \\
 &= \iota_U(\varphi(v))(\beta) && \text{Definition von } \iota_U \\
 &= (\iota_U \circ \varphi)(v)(\beta) && \text{Definition der Komposition von Abb.}
 \end{aligned}$$

Da dies für beliebige $v \in V$ und $\beta \in U^*$ gilt, folgt $\varphi^{tt} \circ \iota_V = \iota_U \circ \varphi$. \square

III.4.14. BEMERKUNG. Die Abbildung $\iota_V: V \rightarrow V^{**}$ ist i.A. nicht surjektiv.

IV. Endlich-dimensionale Vektorräume

Unter einem endlich-dimensionalen Vektorraum verstehen wir einen Vektorraum, der eine endliche Basis besitzt. Die entscheidende Beobachtung ist die Tatsache, dass in diesem Fall je zwei Basen aus gleich vielen Elementen bestehen müssen, siehe Korollar IV.1.5 unten. Dies ermöglicht es jedem endlich-dimensionalen Vektorraum V eine Dimension, $\dim(V) \in \mathbb{N}_0$, zuzuordnen, nämlich die Anzahl der Elemente einer, und dann jeder, Basis von V .

Im ersten Teil dieses Kapitels werden wir die grundlegenden Eigenschaften dieses Dimensionsbegriffs zusammenstellen. Insbesondere werden wir sehen, dass jeder Teilraum W von \mathbb{K}^n endlich-dimensional ist und daher von endlich-vielen linear unabhängigen Vektoren aufgespannt wird. In anderen Worten, jeder Teilraum lässt sich mit Hilfe einer Parameterdarstellung darstellen. Andererseits lässt sich jeder Teilraum W von \mathbb{K}^n auch durch ein homogenes lineares Gleichungssystem beschreiben, wobei mindestens $n - \dim(W)$ viele Gleichungen notwendig sind. Anschließend werden wir uns dem rechnerischen Aspekt widmen und Algorithmen zur Lösung linearer Gleichungssysteme und zur Berechnung der Inversen einer Matrix besprechen. Im letzten Abschnitt werden wir lineare Abbildungen zwischen allgemeinen endlich-dimensionalen Vektorräumen, $\varphi: V \rightarrow W$, durch Matrizen beschreiben indem wir Basen der beiden Vektorräume fixieren.

IV.1. Dimension. Im vorangehenden Kapitel haben wir Basen stets als Teilmengen eines Vektorraums aufgefasst. Manchmal ist es jedoch zweckmäßig *geordnete* Basen zu betrachten. Dies sind Systeme von Vektoren b_1, \dots, b_n die ein linear unabhängiges Erzeugendensystem bilden. Wir wollen damit beginnen diesen Begriff zu präzisieren.

Sei dazu V ein Vektorraum über einem Körper \mathbb{K} . Ein System von Vektoren $v_1, \dots, v_n \in V$ wird *linear abhängig* genannt, falls $\lambda_1 v_1 + \dots + \lambda_n v_n = 0$ für gewisse Skalare $\lambda_1, \dots, \lambda_n \in \mathbb{K}$, die nicht alle verschwinden. Das System v_1, \dots, v_n heißt *linear unabhängig* wenn es nicht linear abhängig ist. In anderen Worten, die Vektoren v_1, \dots, v_n sind genau dann linear unabhängig, wenn aus $\lambda_1 v_1 + \dots + \lambda_n v_n = 0$ stets $\lambda_1 = \dots = \lambda_n = 0$ folgt. Dies ist genau dann der Fall, wenn die Vektoren v_1, \dots, v_n paarweise verschieden sind und die Teilmenge $\{v_1, \dots, v_n\}$ linear unabhängig im Sinn von Abschnitt III.2 ist.

IV.1.1. BEISPIEL. Die Vektoren $\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ sind linear abhängig in \mathbb{R}^2 , aber $\{\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}\}$ bildet eine linear unabhängige Teilmenge von \mathbb{R}^2 .

Ein System von Vektoren $v_1, \dots, v_n \in V$ wird *Erzeugendensystem* von V genannt, falls $\langle v_1, \dots, v_n \rangle = V$ gilt, d.h. falls die Teilmenge $\{v_1, \dots, v_n\}$ ein Erzeugendensystem von V bildet, vgl. Abschnitt III.1. Ein linear unabhängiges Erzeugendensystem v_1, \dots, v_n von V wird als (*geordnete*) *Basis* von V bezeichnet. Die Vektoren v_1, \dots, v_n bilden also genau dann eine geordnete Basis von V , wenn sie paarweise verschieden sind und die Teilmenge $\{v_1, \dots, v_n\}$ eine Basis im Sinn

von Abschnitt III.3 bildet. Aus Proposition III.3.2 erhalten wir sofort folgende Charakterisierung geordneter Basen:

IV.1.2. PROPOSITION. *Für ein System von Vektoren b_1, \dots, b_n eines \mathbb{K} -Vektorraums V sind die folgenden Aussagen äquivalent:*

- (a) *Die Vektoren b_1, \dots, b_n bilden eine Basis von V .*
- (b) *Zu jedem $v \in V$ existieren eindeutige Skalare $\lambda_1, \dots, \lambda_n \in \mathbb{K}$, sodass $v = \lambda_1 b_1 + \dots + \lambda_n b_n$.*
- (c) *Die Abbildung*

$$\phi: \mathbb{K}^n \rightarrow V, \quad \phi \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} := x_1 b_1 + \dots + x_n b_n,$$

ist ein linearer Isomorphismus.

- (d) *Ist W ein \mathbb{K} -Vektorraum und $w_1, \dots, w_n \in W$, dann existiert eine eindeutige lineare Abbildung $\varphi: V \rightarrow W$, sodass $\varphi(b_i) = w_i$ für alle $i = 1, \dots, n$.*

Aus Proposition III.3.14 erhalten wir auch:

IV.1.3. PROPOSITION. *Ist $\varphi: V \rightarrow W$ linear und b_1, \dots, b_n eine Basis von V , dann sind die folgenden Aussagen äquivalent:*

- (a) *φ ist ein Isomorphismus.*
- (b) *Für jede Basis c_1, \dots, c_m von V ist $\varphi(c_1), \dots, \varphi(c_m)$ eine Basis von W .*
- (c) *$\varphi(b_1), \dots, \varphi(b_n)$ ist eine Basis von W .*

Der Schlüssel zum Dimensionsbegriff ist folgendes Resultat.

IV.1.4. SATZ (Austauschsatz von Steinitz). *Sei v_1, \dots, v_n ein Erzeugendensystem eines Vektorraums V , und w_1, \dots, w_k linear unabhängig in V . Dann gilt*

$$k \leq n$$

und, nach geeignetem Umnummerieren der Vektoren v_1, \dots, v_n , bildet auch

$$w_1, \dots, w_k, v_{k+1}, \dots, v_n$$

ein Erzeugendensystem von V .

BEWEIS. Wir führen den Beweis mittels Induktion nach k . Für $k = 0$ ist die Aussage trivial. Für den Induktionsschritt sei nun $k \geq 1$ und die Aussage für $k - 1$ bereits gezeigt. Nach Induktionsvoraussetzung gilt daher $k - 1 \leq n$ und, nach geeignetem Umnummerieren der Vektoren v_1, \dots, v_n bildet

$$w_1, \dots, w_{k-1}, v_k, \dots, v_n \tag{IV.1}$$

ein Erzeugendensystem von V . Daher existieren Skalare $\lambda_i \in \mathbb{K}$, sodass

$$w_k = \lambda_1 w_1 + \dots + \lambda_{k-1} w_{k-1} + \lambda_k v_k + \dots + \lambda_n v_n. \tag{IV.2}$$

Da das System w_1, \dots, w_k linear unabhängig ist folgt $k \leq n$ und mindestens einer der Skalare $\lambda_k, \dots, \lambda_n$ muss verschieden von 0 sein. Andernfalls erhielten

wir die Relation $w_k = \lambda_1 w_1 + \dots + \lambda_{k-1} w_{k-1}$, was der linearen Unabhängigkeit des Systems w_1, \dots, w_k widerspräche. Durch Umm Nummerieren der Vektoren v_i können wir also $\lambda_k \neq 0$ erreichen. Aus (IV.2) erhalten wir daher

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

Dies zeigt, dass der Vektor v_k im Erzeugniss des Systems

$$w_1, \dots, w_k, v_{k+1}, \dots, v_n \tag{IV.3}$$

liegt. Mit (IV.1) ist daher auch (IV.3) ein Erzeugendensystem für V . \square

IV.1.5. KOROLLAR. *Für einen Vektorraum V sind äquivalent:*

- (a) V besitzt eine endliche Basis.
- (b) V ist endlich erzeugt, d.h. besitzt ein endliches Erzeugendensystem.
- (c) Jede linear unabhängige Teilmenge von V ist endlich.
- (d) Jede Basis von V ist endlich.

In diesem Fall haben je zwei Basen von V gleich viele Elementen.

BEWEIS. Die Implikation (a) \Rightarrow (b) ist trivial. Die Implikation (b) \Rightarrow (c) folgt aus Satz IV.1.4. Die Implikation (c) \Rightarrow (d) ist trivial. Die Implikation (d) \Rightarrow (a) folgt aus Korollar III.3.23(a). Damit ist die Äquivalenz der vier Aussagen gezeigt. Sind b_1, \dots, b_m und b'_1, \dots, b'_n zwei endliche Basen von V , dann erhalten wir aus Satz IV.1.4 nun $m \leq n$ und $n \leq m$, also $n = m$, d.h. die Basen bestehen aus gleich vielen Vektoren. \square

IV.1.6. DEFINITION (Dimension). Ein Vektorraum V wird *endlich dimensional* genannt, wenn er eine endliche Basis besitzt. In diesem Fall existiert eine eindeutige Zahl $n \in \mathbb{N}_0$, sodass jede Basis von V aus genau n Elementen besteht, siehe Korollar IV.1.5. Diese Zahl n wird als *Dimension* von V bezeichnet und mit $\dim(V)$ notiert. Wir sagen auch V ist ein n -dimensionaler Vektorraum. Besitzt V keine endliche Basis dann wird V *unendlich dimensional* genannt und wir schreiben $\dim(V) = \infty$.

IV.1.7. BEMERKUNG. Ein Vektorraum V ist genau dann 0-dimensional, wenn die leere Menge eine Basis von V bildet. Dies ist genau dann der Fall, wenn V nur aus dem Nullvektor besteht, d.h. $V = \{0\}$.

IV.1.8. BEISPIEL. Es gilt $\dim(\mathbb{K}^n) = n$, denn nach Beispiel III.3.5 bilden die Einheitsvektoren e_1, \dots, e_n eine Basis von \mathbb{K}^n , die aus genau n Vektoren besteht.

IV.1.9. BEISPIEL. Es ist $\dim(\mathbb{K}[z]_{\leq n}) = n+1$, denn die Monome $1, z, z^2, \dots, z^n$ bilden eine Basis von $\mathbb{K}[z]_{\leq n}$, die aus genau $n+1$ Elementen besteht.

IV.1.10. BEISPIEL. Es gilt $\dim(M_{m \times n}(\mathbb{K})) = mn$, denn die Matrizen $E_{i,j}$, $1 \leq i \leq m$, $1 \leq j \leq n$, bilden eine Basis von $M_{m \times n}(\mathbb{K})$, die aus genau mn vielen

Elementen besteht. Dabei bezeichnet

$$E_{i,j} = \begin{pmatrix} 0 & \cdots & 0 & 0 & 0 & \cdots & 0 \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ 0 & \cdots & 0 & 0 & 0 & \cdots & 0 \\ 0 & \cdots & 0 & 1 & 0 & \cdots & 0 \\ 0 & \cdots & 0 & 0 & 0 & \cdots & 0 \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ 0 & \cdots & 0 & 0 & 0 & \cdots & 0 \end{pmatrix}$$

jene $(m \times n)$ -Matrix, deren einzige nicht verschwindende Eintragungen eine Eins in der i -ten Zeile und j -ten Spalte bildet. In anderen Worten: $(E_{i,j})_{kl} = \delta_{ik}\delta_{jl}$.

IV.1.11. BEISPIEL. Der Vektorraum der Polynome, $\mathbb{K}[z]$, ist unendlich-dimensional, denn die linear unabhängige Teilmenge der Monome, $\{1, z, z^2, z^3, \dots\}$, hat unendlich viele Elemente.

IV.1.12. BEISPIEL. Betrachten wir \mathbb{C} als komplexen Vektorraum, so ist dieser ein-dimensional, denn 1 bildet eine Basis. Wir können $\mathbb{C} = \mathbb{R}^2$ aber auch als zwei-dimensionalen reellen Vektorraum auffassen, dann bildet etwa $1, \mathbf{i}$ eine Basis. Es gilt daher $\dim_{\mathbb{C}}(\mathbb{C}) = 1$ und $\dim_{\mathbb{R}}(\mathbb{C}) = 2$, siehe auch Aufgabe 75.

IV.1.13. BEMERKUNG. Sind V und W zwei isomorphe Vektorräume und ist V endlich dimensional, dann ist auch W endlich-dimensional und es gilt $\dim(V) = \dim(W)$. Ist nämlich b_1, \dots, b_n eine Basis von V und $\varphi: V \rightarrow W$ ein Isomorphismus, dann bildet $\varphi(b_1), \dots, \varphi(b_n)$ eine Basis von W , siehe Proposition IV.1.3.

Das folgende Resultat zeigt, dass es über jedem Körper \mathbb{K} im Wesentlichen nur einen n -dimensionalen Vektorraum gibt, nämlich \mathbb{K}^n .

IV.1.14. KOROLLAR. *Zwei endlich dimensionale \mathbb{K} -Vektorräume sind genau dann isomorph, wenn sie gleiche Dimension haben. Insbesondere gilt $\mathbb{K}^n \cong \mathbb{K}^m$ genau dann, wenn $n = m$. Jeder endlich-dimensionale \mathbb{K} -Vektorraum V ist zu \mathbb{K}^n isomorph, wobei $n = \dim(V)$.*

BEWEIS. In Bemerkung IV.1.13 haben wir bereits beobachtet, dass isomorphe endlich-dimensionale Vektorräume gleiche Dimension haben müssen. Umgekehrt folgt aus Bemerkung III.3.11, dass endlich-dimensionale Vektorräume gleicher Dimension isomorph sind. \square

IV.1.15. BEMERKUNG. Sind $A \in M_{m \times n}(\mathbb{K})$ und $B \in M_{n \times m}(\mathbb{K})$ zwei Matrizen, sodass $AB = I_m$ und $BA = I_n$, dann muss $n = m$ gelten. Eine Matrix kann also nur dann invertierbar sein, wenn sie quadratisch ist. Aus den beiden Gleichungen folgt nämlich, dass die assoziierten linearen Abbildungen $\psi_A: \mathbb{K}^n \rightarrow \mathbb{K}^m$ und $\psi_B: \mathbb{K}^m \rightarrow \mathbb{K}^n$ zueinander inverse Isomorphismen darstellen, d.h. $\psi_A \circ \psi_B = \psi_{AB} = \psi_{I_m} = \text{id}_{\mathbb{K}^m}$ und $\psi_B \circ \psi_A = \psi_{BA} = \psi_{I_n} = \text{id}_{\mathbb{K}^n}$, vgl. Satz II.4.4, die Relation $n = m$ folgt daher aus Korollar IV.1.14.

IV.1.16. KOROLLAR. *Ist V ein n -dimensionaler Vektorraum dann gilt:*

- (a) *Jede linear unabhängige Teilmenge von V hat höchstens n verschiedene Elemente. Sind v_1, \dots, v_n linear unabhängig in V , dann bilden diese Vektoren schon eine Basis von V .*
- (b) *Jedes Erzeugendensystem von V besitzt mindestens n verschiedene Elemente. Ist v_1, \dots, v_n ein Erzeugendensystem von V , dann bilden diese Vektoren schon eine Basis von V .*

BEWEIS. Nach Voraussetzung existiert ein linear unabhängiges Erzeugendensystem von V , das aus genau n Vektoren besteht. Nach Satz IV.1.4 kann eine linear unabhängige Teilmenge von V höchstens n Elementen haben, und jedes Erzeugendensystem von V muss aus mindestens n Vektoren bestehen. Die restlichen Behauptungen folgen aus Proposition III.3.19. \square

Daraus erhalten wir auch:

IV.1.17. KOROLLAR. *Sei V ein n -dimensionaler Vektorraum und v_1, \dots, v_n ein System von n Vektoren in V . Dann sind äquivalent:*

- (a) *v_1, \dots, v_n ist eine Basis von V .*
- (b) *v_1, \dots, v_n ist linear unabhängig in V .*
- (c) *v_1, \dots, v_n ist ein Erzeugendensystem von V .*

IV.1.18. BEMERKUNG. Nach Bemerkung III.3.15 ist eine quadratische Matrix $A \in M_{n \times n}(\mathbb{K})$ genau dann invertierbar, wenn ihre Spaltenvektoren eine Basis von \mathbb{K}^n bilden. Nach Korollar IV.1.17 ist dies genau dann der Fall, wenn die Spaltenvektoren linear unabhängig sind. Ebenso ist A genau dann invertierbar, wenn die ihre Spaltenvektoren ein Erzeugendensystem von \mathbb{K}^n bilden.

IV.1.19. BEMERKUNG. Ein Vektorraum V hat genau dann Dimension n , wenn es n linear unabhängige Vektoren v_1, \dots, v_n in V gibt und je $n+1$ Vektoren in V linear abhängig sind. Auch hat V genau dann Dimension n , wenn ein Erzeugendensystem von V mit n Vektoren existiert, und je $n-1$ Vektoren V nicht erzeugen. Dies folgt aus Korollar IV.1.16 und Proposition III.3.19, vgl. Aufgabe 67.

IV.1.20. KOROLLAR. *Ist W ein Teilraum eines endlich-dimensionalen Vektorraums V , dann ist auch W endlich-dimensional und es gilt*

$$\dim(W) \leq \dim(V).$$

Ist darüber hinaus $\dim(W) = \dim(V)$, dann folgt schon $W = V$.

BEWEIS. Nach Korollar IV.1.16(a) besteht jede linear unabhängige Teilmenge von W aus höchstens $\dim(V)$ vielen Elementen. Es gibt daher eine größte Zahl $k \in \mathbb{N}_0$, sodass linear unabhängige Vektoren w_1, \dots, w_k in W existieren. Nach Konstruktion bilden die Vektoren w_1, \dots, w_k eine maximal linear unabhängige Teilmenge von W und daher eine Basis von W , siehe Proposition III.3.19. Somit ist W endlich dimensional, und es gilt $\dim(W) = k \leq \dim(V)$. Gilt darüber

hinaus $\dim(W) = \dim(V)$, dann ist w_1, \dots, w_k auch Basis von V , siehe Korollar IV.1.16(a), und daher $V = \langle w_1, \dots, w_k \rangle = W$. \square

IV.1.21. BEMERKUNG (Teilräume von \mathbb{K}^n). Nach Korollar IV.1.20 ist jeder Teilraum $W \subseteq \mathbb{K}^n$ endlich-dimensional und es gilt $0 \leq \dim(W) \leq n$. Es existieren daher linear unabhängige Vektoren $w_1, \dots, w_m \in W$, sodass $W = \langle w_1, \dots, w_m \rangle$, wobei $m = \dim(W)$. Diese Vektoren bilden eine Basis von W und es gilt $W \cong \mathbb{K}^m$. Jeder Teilraum von \mathbb{K}^n ist also zu einem \mathbb{K}^m isomorph, wobei $0 \leq m \leq n$.

IV.1.22. BEISPIEL (Teilräume von \mathbb{K}^2). Aus Bemerkung IV.1.21 folgt, dass jeder Teilraum W von \mathbb{K}^2 von der Form

$$W = \{0\}, \quad W = \langle w \rangle, \quad \text{oder} \quad W = \mathbb{K}^2$$

ist, wobei $0 \neq w \in \mathbb{K}^2$. In Beispiel II.2.9 haben wir dies schon auf elementare Weise hergeleitet.

IV.1.23. BEISPIEL (Teilräume von \mathbb{K}^3). Bemerkung IV.1.21 folgt, dass jeder Teilraum W von \mathbb{K}^3 von der Form

$$W = \{0\}, \quad W = \langle w \rangle, \quad W = \langle w_1, w_2 \rangle, \quad \text{oder} \quad W = \mathbb{K}^3$$

ist, wobei $0 \neq w \in \mathbb{K}^3$ und w_1, w_2 linear unabhängig in \mathbb{K}^3 sind. Für $\mathbb{K} = \mathbb{R}$ entsprechen die nicht-trivialen Fälle also genau den Geraden bzw. Ebenen durch den Koordinatenursprung.

IV.2. Dimensionsformeln. Wir wollen in diesem Abschnitt Dimensionsformel für Summen und Durchschnitte von Teilräumen, Kern und Bild linearer Abbildungen, Dualräumen, Quotientenräumen und Annihilatoren herleiten, und einige Anwendungen besprechen.

IV.2.1. SATZ (Dimension von Summen und Durchschnitten). *Seien W_1 und W_2 zwei Teilräume eines Vektorraums V . Es ist $W_1 + W_2$ genau dann endlich-dimensional, wenn W_1 und W_2 beide endlich-dimensional sind. In diesem Fall ist auch $W_1 \cap W_2$ endlich-dimensional und es gilt*

$$\dim(W_1 \cap W_2) + \dim(W_1 + W_2) = \dim(W_1) + \dim(W_2).$$

BEWEIS. Ist $W_1 + W_2$ endlich-dimensional, dann gilt dies auch für die Teilräume W_1 , W_2 und $W_1 \cap W_2$ von $W_1 + W_2$, siehe Korollar IV.1.20. Seien nun umgekehrt W_1 und W_2 endlich-dimensional. Nach Proposition III.3.29 existieren daher endliche Basen B_1 von W_1 und B_2 von W_2 , sodass $B_1 \cup B_2$ eine Basis von $W_1 + W_2$ bildet und $B_1 \cap B_2$ eine Basis von $W_1 \cap W_2$ ist. Es gilt daher:

$$\begin{aligned} \dim(W_1) &= \#B_1 \\ \dim(W_2) &= \#B_2 \\ \dim(W_1 \cap W_2) &= \#(B_1 \cap B_2) \\ \dim(W_1 + W_2) &= \#(B_1 \cup B_2) \end{aligned}$$

wobei $\#X$ die Anzahl der Elemente einer endlichen Menge X bezeichnet. Aus der evidenten Formel

$$\#(B_1 \cap B_2) + \#(B_1 \cup B_2) = \#B_1 + \#B_2$$

erhalten wir den Satz. □

IV.2.2. KOROLLAR (Dimension direkter Summen). *Seien W_1 und W_2 zwei komplementäre Teilräume eines Vektorraums V , d.h. $V = W_1 \oplus W_2$. Es ist V genau dann endlich-dimensional, wenn W_1 und W_2 beide endlich-dimensional sind, und in diesem Fall gilt*

$$\dim(V) = \dim(W_1) + \dim(W_2).$$

BEWEIS. Nach Voraussetzung gilt $W_1 \cap W_2 = \{0\}$ und $W_1 + W_2 = V$. Das Korollar folgt daher sofort aus Satz IV.2.1. □

IV.2.3. KOROLLAR. *Seien W_1 und W_2 zwei Teilräume eines endlich-dimensionalen Vektorraums V und $\dim(W_1) + \dim(W_2) = \dim(V)$. Dann sind die folgenden Aussagen äquivalent:*

- (a) $W_1 \oplus W_2 = V$.
- (b) $W_1 \cap W_2 = \{0\}$.
- (c) $W_1 + W_2 = V$

BEWEIS. Nach Satz IV.2.1 ist $\dim(W_1 \cap W_2) + \dim(W_1 + W_2) = \dim(V)$, also

$$\dim(W_1 \cap W_2) = 0 \quad \Leftrightarrow \quad \dim(W_1 + W_2) = \dim(V).$$

Mit Korollar IV.1.20 und Bemerkung IV.1.7 folgt daher die Äquivalenz (b) \Leftrightarrow (c). Die verbleibenden Behauptungen sind nun trivial. □

IV.2.4. BEISPIEL. Ist E ein 2-dimensionaler Teilraum von \mathbb{K}^3 und ist L ein 1-dimensionaler Teilraum von \mathbb{K}^3 , der nicht in E enthalten ist, dann gilt schon

$$E \oplus L = \mathbb{K}^3.$$

Nach Voraussetzung ist nämlich $E \cap L \subsetneq L$, also $\dim(E \cap L) < \dim(L) = 1$ nach Korollar IV.1.20, folglich $\dim(E \cap L) = 0$ und daher $E \cap L = \{0\}$. Die Behauptung folgt somit aus Korollar IV.2.3.

IV.2.5. BEMERKUNG. Sind W_1 und W_2 zwei Teilräume eines endlich-dimensionalen Vektorraums V , sodass

$$\dim(W_1 \cap W_2) \leq \dim(W_1) + \dim(W_2) - \dim(V),$$

dann muss schon $W_1 + W_2 = V$ und

$$\dim(W_1 \cap W_2) = \dim(W_1) + \dim(W_2) - \dim(V)$$

gelten. Aus Satz IV.2.1 folgt nämlich

$$\dim(W_1 + W_2) = \dim(W_1) + \dim(W_2) - \dim(W_1 \cap W_2) \geq \dim(V),$$

also $W_1 + W_2 = V$ nach Korollar IV.1.20. Die Formel für die Dimension des Durchschnitts folgt nun aus Satz IV.2.1.

IV.2.6. BEISPIEL. Sind E_1 und E_2 zwei verschiedene 2-dimensionale Teilräume von \mathbb{K}^3 , dann gilt

$$E_1 + E_2 = \mathbb{K}^3 \quad \text{und} \quad \dim(E_1 \cap E_2) = 1.$$

Nach Voraussetzung ist nämlich $E_1 \cap E_2 \subsetneq E_1$ oder $E_1 \cap E_2 \subsetneq E_2$, jedenfalls folgt $\dim(E_1 \cap E_2) < 2 = \dim(E_1) = \dim(E_2)$ nach Korollar IV.1.20, und somit $\dim(E_1 \cap E_2) \leq 1 = 2 + 2 - 3 = \dim(E_1) + \dim(E_2) - \dim(\mathbb{K}^3)$. Die Behauptung folgt daher aus Bemerkung IV.2.5. Für den Körper $\mathbb{K} = \mathbb{R}$ bedeutet dies, dass zwei verschiedene Ebenen durch den Koordinatenursprung ganz \mathbb{R}^3 erzeugen und ihr Durchschnitt eine Gerade bildet. Siehe auch Aufgabe 71.

IV.2.7. BEMERKUNG. Sind W_1 und W_2 zwei Teilräume eines endlich-dimensionalen Vektorraums V , sodass

$$\dim(W_1 + W_2) \geq \dim(W_1) + \dim(W_2),$$

dann muss schon $W_1 \cap W_2 = \{0\}$ und

$$\dim(W_1 + W_2) = \dim(W_1) + \dim(W_2)$$

gelten. Aus Satz IV.2.1 folgt nämlich

$$\dim(W_1 \cap W_2) = \dim(W_1) + \dim(W_2) - \dim(W_1 + W_2) \leq 0,$$

also $\dim(W_1 \cap W_2) = 0$ und daher $W_1 \cap W_2 = \{0\}$. Die Formel für die Dimension der Summe folgt nun aus Satz IV.2.1.

IV.2.8. KOROLLAR (Dimension eines Quotienten). *Sei W ein Teilraum eines Vektorraums V . Es ist V genau dann endlich-dimensional, wenn W und V/W beide endlich-dimensional sind, und in diesem Fall gilt*

$$\dim(V) = \dim(W) + \dim(V/W).$$

BEWEIS. Nach Korollar III.3.26 existiert ein zu W komplementärer Teilraum W' in V , d.h. $V = W \oplus W'$. Nach Korollar II.6.7 gilt $W' \cong V/W$. Insbesondere ist W' genau dann endlich-dimensional, wenn V/W endlich-dimensional ist, und in diesem Fall gilt $\dim(W') = \dim(V/W)$. Das Korollar folgt daher aus Korollar IV.2.2. \square

IV.2.9. KOROLLAR (Dimension von Kern und Bild). *Sei $\varphi: V \rightarrow W$ eine lineare Abbildung. Es ist V genau dann endlich-dimensional, wenn $\ker(\varphi)$ und $\text{img}(\varphi)$ beide endlich-dimensional sind, und in diesem Fall gilt*

$$\dim(V) = \dim(\ker(\varphi)) + \dim(\text{img}(\varphi)).$$

BEWEIS. Nach Korollar II.6.6 gilt $V/\ker(\varphi) \cong \text{img}(\varphi)$. Insbesondere ist also $\text{img}(\varphi)$ genau dann endlich-dimensional, wenn $V/\ker(\varphi)$ endlich-dimensional ist, und in diesem Fall gilt $\dim(V/\ker(\varphi)) = \dim(\text{img}(\varphi))$. Das Korollar folgt daher aus Korollar IV.2.8. \square

IV.2.10. BEISPIEL (Hyperebenen). Sei V ein n -dimensionaler \mathbb{K} -Vektorraum und $\alpha: V \rightarrow \mathbb{K}$ ein nicht-triviales lineares Funktional, $\alpha \neq 0$. Dann folgt $\{0\} \subsetneq \text{img}(\alpha) \subseteq \mathbb{K}$, aus Dimensionsgründen gilt daher $\text{img}(\alpha) = \mathbb{K}$, also $\dim(\text{img}(\alpha)) = 1$. Mit Korollar IV.2.9 folgt $\dim(\ker(\alpha)) = n - 1$. Teilräume der Form $\ker(\alpha)$ mit $0 \neq \alpha \in V^*$ werden *Hyperebenen* genannt. Dies sind also genau jene Teilräume, die sich durch *eine* nicht-triviale lineare Gleichung beschreiben lassen. In \mathbb{R}^3 entsprechen diese genau den Ebenen durch den Koordinatenursprung.

IV.2.11. KOROLLAR. *Ist $\varphi: V \rightarrow W$ eine lineare Abbildung zwischen endlich-dimensionalen \mathbb{K} -Vektorräumen gleicher Dimension, $\dim(V) = \dim(W)$, dann sind die folgenden Aussagen äquivalent:*

- (a) φ ist injektiv.
- (b) φ ist surjektiv.
- (c) φ ist ein linearer Isomorphismus.

BEWEIS. Es gilt:

$$\begin{array}{ll}
 \varphi \text{ ist injektiv} \Leftrightarrow \ker(\varphi) = \{0\} & \text{nach Proposition II.3.22} \\
 \Leftrightarrow \dim(\ker(\varphi)) = 0 & \text{nach Bemerkung IV.1.7} \\
 \Leftrightarrow \dim(\text{img}(\varphi)) = \dim(V) & \text{nach Korollar IV.2.9} \\
 \Leftrightarrow \dim(\text{img}(\varphi)) = \dim(W) & \text{da } \dim(V) = \dim(W) \\
 \Leftrightarrow \text{img}(\varphi) = W & \text{nach Korollar IV.1.20} \\
 \Leftrightarrow \varphi \text{ ist surjektiv} &
 \end{array}$$

Dies zeigt (a) \Leftrightarrow (b). Die verbleibenden Behauptungen folgen sofort aus Bemerkung II.3.11. \square

IV.2.12. BEMERKUNG. Eine injektive lineare Abbildung zwischen unendlich-dimensionalen Vektorräumen wird i.A. nicht surjektiv sein. Etwa ist die lineare Abbildung $\mathbb{K}[z] \rightarrow \mathbb{K}[z]$, $p \mapsto zp$, injektiv aber nicht surjektiv. Auch muss eine surjektive lineare Abbildung zwischen unendlich-dimensionalen Vektorräumen nicht injektiv sein. Zum Beispiel ist die lineare Abbildung $\mathbb{K}[z] \rightarrow \mathbb{K}[z]$, $a_0 + a_1z + a_2z^2 + a_3z^3 + \dots \mapsto a_1 + a_2z + a_3z^2 + \dots$, surjektiv aber nicht injektiv. Auf die Voraussetzung in Korollar IV.2.11 kann daher nicht verzichtet werden.

IV.2.13. KOROLLAR. *Seien $A, A' \in M_{n \times n}(\mathbb{K})$ zwei quadratische Matrizen.*

- (a) *Gilt $A'A = I_n$, dann ist A invertierbar mit Inverser $A^{-1} = A'$.*
- (b) *Gilt $AA' = I_n$, dann ist A invertierbar mit Inverser $A^{-1} = A'$.*

BEWEIS. Wir zeigen nur die erste Behauptung, die zweite lässt sich völlig analog beweisen. Seien also $A, A' \in M_{n \times n}(\mathbb{K})$ und $A'A = I_n$. Betrachten wir die assoziierten linearen Abbildungen $\psi_A, \psi_{A'}: \mathbb{K}^n \rightarrow \mathbb{K}^n$, dann folgt $\text{id}_{\mathbb{K}^n} = \psi_{I_n} = \psi_{A'A} = \psi_{A'} \circ \psi_A$, siehe Satz II.4.4. Die lineare Abbildung $\psi_{A'}: \mathbb{K}^n \rightarrow \mathbb{K}^n$ ist daher injektiv, und somit ein Isomorphismus, siehe Korollar IV.2.11. Nach Korollar II.4.6 ist A also invertierbar. \square

IV.2.14. SATZ (Dimension von $L(V, W)$). *Sind V und W zwei endlich-dimensionale \mathbb{K} -Vektorräume, dann ist auch $L(V, W)$ endlich-dimensional und es gilt*

$$\dim(L(V, W)) = \dim(V) \cdot \dim(W).$$

PROOF. Es existieren Isomorphismen $\phi: \mathbb{K}^n \xrightarrow{\cong} V$ und $\psi: \mathbb{K}^m \xrightarrow{\cong} W$, wobei $n = \dim(V)$ und $m = \dim(W)$. Weiters ist die Zuordnung

$$L(V, W) \xrightarrow{\cong} L(\mathbb{K}^n, \mathbb{K}^m), \quad \rho \mapsto \psi^{-1} \circ \rho \circ \phi,$$

ein linearer Isomorphismus mit Umkehrabbildung $\sigma \mapsto \psi \circ \sigma \circ \phi^{-1}$. Die Linearität dieser Abbildung folgt aus Proposition II.3.18, denn

$$\psi^{-1} \circ (\rho_1 + \rho_2) \circ \phi = \psi^{-1} \circ \rho_1 \circ \phi + \psi^{-1} \circ \rho_2 \circ \phi$$

und $\psi^{-1} \circ (\lambda\rho) \circ \phi = \lambda(\psi^{-1} \circ \rho \circ \phi)$, für beliebige $\rho, \rho_1, \rho_2 \in L(V, W)$ und $\lambda \in \mathbb{K}$. Da $L(\mathbb{K}^n, \mathbb{K}^m) \cong M_{m \times n}(\mathbb{K})$, siehe Satz II.4.4, erhalten wir $L(V, W) \cong M_{m \times n}(\mathbb{K})$, das Korollar folgt somit aus Beispiel IV.1.10, $\dim(M_{m \times n}(\mathbb{K})) = mn$. \square

IV.2.15. KOROLLAR. *Ein Vektorraum V ist genau dann endlich-dimensional, wenn sein Dualraum V^* endlich-dimensional ist. In diesem Fall gilt*

$$\dim(V) = \dim(V^*)$$

und die kanonische Abbildung $\iota: V \rightarrow V^{**}$ aus Proposition III.4.13 ist ein Isomorphismus.

BEWEIS. Ist V endlich-dimensional, dann folgt mit Satz IV.2.14, dass auch $V^* = L(V, \mathbb{K})$ endlich-dimensional ist und

$$\dim(V^*) = \dim(L(V, \mathbb{K})) = \dim(V) \cdot \dim(\mathbb{K}) = \dim(V) \cdot 1 = \dim(V).$$

Nach Proposition III.4.13 ist die Abbildung $\iota: V \rightarrow V^{**}$ injektiv, aus Korollar IV.2.11 folgt daher, dass ι ein Isomorphismus ist, denn $\dim(V^{**}) = \dim(V^*) = \dim(V)$. Sei nun umgekehrt V^* endlich dimensional. Nach dem eben gezeigten ist daher auch V^{**} endlich-dimensional. Nach Proposition III.4.13 ist die kanonische lineare Abbildung $\iota: V \rightarrow V^{**}$ injektiv. Somit ist V zu dem Teilraum $\text{img}(\iota)$ von V^{**} isomorph. Aus Korollar IV.1.20 folgt daher, dass auch V endlich-dimensional sein muss. \square

IV.2.16. KOROLLAR (Dimension des Annihilators). *Sei W ein Teilraum eines Vektorraums V . Es ist W° genau dann endlich-dimensional, wenn V/W endlich-dimensional ist, und in diesem Fall gilt*

$$\dim(W^\circ) = \dim(V/W).$$

BEWEIS. Dies folgt aus Korollar IV.2.15, denn es gilt $W^\circ \cong (V/W)^*$, siehe Korollar III.4.12(b). \square

IV.2.17. SATZ (Teilräume und Gleichungssysteme). Sei W ein Teilraum eines Vektorraums V , sodass V/W endlich-dimensional ist. Eine Teilmenge $A \subseteq V^*$ ist genau dann ein Erzeugendensystem von W° , wenn $W = \bigcap_{\alpha \in A} \ker(\alpha)$ gilt. Es gibt daher $k = \dim(V/W)$ viele Funktionale $\alpha_1, \dots, \alpha_k \in V^*$, sodass

$$W = \bigcap_{i=1}^k \ker(\alpha_i),$$

d.h. W ist Durchschnitt von k Hyperebenen, vgl. Beispiel IV.2.10. Jedes Funktional $\alpha \in V^*$, das auf W verschwindet, $\alpha|_W = 0$, ist eine Linearkombination von $\alpha_1, \dots, \alpha_k$. Mit weniger als k linearen Funktionalen lässt sich W nicht beschreiben.

BEWEIS. In Bemerkung III.4.8 haben wir bereits festgehalten, dass für jedes Erzeugendensystem A von W° auch $W = \bigcap_{\alpha \in A} \ker(\alpha)$ gelten muss. Sei nun umgekehrt $W = \bigcap_{\alpha \in A} \ker(\alpha)$. Nach Korollar IV.2.16 ist W° endlich-dimensional, also ist auch der Teilraum $\langle A \rangle$ endlich-dimensional, es existieren daher endlich viele Funktionale $\alpha_1, \dots, \alpha_l \in A$, die eine Basis von $\langle A \rangle$ bilden. Wie in Bemerkung III.4.8 folgt

$$W = \bigcap_{\alpha \in A} \ker(\alpha) = \bigcap_{\alpha \in \langle A \rangle} \ker(\alpha) = \bigcap_{i=1}^l \ker(\alpha_i),$$

denn $\langle A \rangle = \langle \alpha_1, \dots, \alpha_l \rangle$. Zusammen definieren die Funktionale $\alpha_i: V \rightarrow \mathbb{K}$ eine lineare Abbildung $\varphi: V \rightarrow \mathbb{K}^l$, sodass $W = \ker(\varphi)$. Wir erhalten $V/W \cong \text{img}(\varphi) \subseteq \mathbb{K}^l$, also

$$\dim(W^\circ) = \dim(V/W) = \dim(\text{img}(\varphi)) \leq \dim(\mathbb{K}^l) = l.$$

Aus Korollar IV.1.16(a) folgt daher, dass $\alpha_1, \dots, \alpha_l$ eine Basis von W° sein muss, denn $\alpha_1, \dots, \alpha_l$ sind linear unabhängig in W° . Somit enthält A ein Erzeugendensystem von W° und muss daher selbst ein Erzeugendensystem von W° sein. Die verbleibenden Aussagen folgen nun aus Korollar IV.2.16. \square

IV.2.18. BEMERKUNG (Teilräume von \mathbb{K}^n und Gleichungssysteme). Sei W ein Teilraum von \mathbb{K}^n und $k = n - \dim(W)$. Dann existiert ein Gleichungssystem

$$\begin{aligned} a_{11}x_1 + \cdots + a_{1n}x_n &= 0 \\ &\vdots \\ a_{k1}x_1 + \cdots + a_{kn}x_n &= 0 \end{aligned}$$

sodass W mit der Lösungsmenge dieses Systems übereinstimmt. Dies folgt aus Satz IV.2.17, denn $\dim(\mathbb{K}^n/W) = \dim(\mathbb{K}^n) - \dim(W) = k$. Jeder Teilraum von \mathbb{K}^n lässt sich daher durch ein Gleichungssystem beschreiben. Fassen wir die Koeffizienten a_{ij} zu einer Matrix $A \in M_{k \times n}(\mathbb{K})$ zusammen, dann gilt also $W = \{x \in \mathbb{K}^n : Ax = 0\}$. Weniger als k Gleichungen reichen nicht aus um W darzustellen, auch dies folgt aus Satz IV.2.17.

IV.2.19. DEFINITION (Rang linearer Abbildungen). Eine lineare Abbildung $\varphi: V \rightarrow W$ hat *endlichen Rang* wenn ihr Bild, $\text{img}(\varphi)$, endlich-dimensional ist. In diesem Fall wird die Zahl

$$\text{rank}(\varphi) := \dim(\text{img}(\varphi))$$

als *Rang* der linearen Abbildung φ bezeichnet.

IV.2.20. SATZ (Rang der dualen Abbildung). *Eine lineare Abbildung $\varphi: V \rightarrow W$ hat genau dann endlichen Rang, wenn ihre duale Abbildung $\varphi^t: W^* \rightarrow V^*$ endlichen Rang hat, und in diesem Fall gilt*

$$\text{rank}(\varphi^t) = \text{rank}(\varphi).$$

BEWEIS. Aus Satz III.4.10, Korollar III.4.12(b) und Korollar II.6.6 erhalten wir Isomorphismen

$$\text{img}(\varphi^t) = \ker(\varphi)^\circ \cong (V/\ker(\varphi))^* \cong \text{img}(\varphi)^*.$$

Insbesondere ist $\text{img}(\varphi^t)$ genau dann endlich-dimensional, wenn $\text{img}(\varphi)^*$ endlich-dimensional ist, und in diesem Fall gilt $\dim(\text{img}(\varphi^t)) = \dim(\text{img}(\varphi)^*)$. Der Satz folgt daher aus Korollar IV.2.15. \square

IV.2.21. BEMERKUNG. Jeder komplexe Vektorraum V kann auch als reeller Vektorraum aufgefasst werden, indem wir die Skalarmultiplikation $\mathbb{C} \times V \rightarrow V$ zu $\mathbb{R} \times V \rightarrow V$ einschränken. Wir werden diesen reellen Vektorraum mit $V^{\mathbb{R}}$ bezeichnen, er wird der dem komplexen Vektorraum V zugrundeliegende reelle Vektorraum genannt. Ist b_1, \dots, b_n eine Basis des komplexen Vektorraums V , dann bildet

$$b_1, \mathbf{i}b_1, b_2, \mathbf{i}b_2, \dots, b_n, \mathbf{i}b_n$$

eine Basis des zugrundeliegenden reellen Vektorraums $V^{\mathbb{R}}$, vgl. Aufgabe 76. Mit V ist daher auch $V^{\mathbb{R}}$ endlich-dimensional und es gilt

$$\dim_{\mathbb{R}}(V^{\mathbb{R}}) = 2 \dim_{\mathbb{C}}(V).$$

IV.2.22. BEMERKUNG. Sei V ein reeller Vektorraum. Mit den Operationen

$$(V \times V) \times (V \times V) \xrightarrow{+} V \times V, \quad (v_1, w_1) + (v_2, w_2) := (v_1 + v_2, w_1 + w_2),$$

$$\mathbb{C} \times (V \times V) \xrightarrow{\cdot} V \times V, \quad (a + \mathbf{b}\mathbf{i})(v, w) := (av - bw, bv + aw),$$

$a, b \in \mathbb{R}$, $v, v_1, v_2, w, w_1, w_2 \in V$, wird $V \times V$ zu einem komplexen Vektorraum, vgl. Aufgabe 77. Wir bezeichnen diesen komplexen Vektorraum mit $V^{\mathbb{C}}$, er wird die *Komplexifizierung* von V genannt. Wir fassen V als Teilmenge von $V^{\mathbb{C}}$ auf, $v \mapsto (v, 0)$. Jedes Element von $V^{\mathbb{C}}$ lässt sich daher in der Form $v + \mathbf{i}w$ schreiben, für eindeutig bestimmte $v, w \in V$. In dieser Darstellung sehen Addition und Skalarmultiplikation vertrauter aus:

$$(v_1 + \mathbf{i}w_1) + (v_2 + \mathbf{i}w_2) = (v_1 + v_2) + \mathbf{i}(w_1 + w_2)$$

$$(a + \mathbf{b}\mathbf{i})(v + \mathbf{i}w) = av - bw + \mathbf{i}(aw + bv)$$

Ist b_1, \dots, b_n eine Basis des reellen Vektorraums V , dann bildet b_1, \dots, b_n auch eine Basis des komplexen Vektorraums $V^{\mathbb{C}}$. Mit V ist daher auch $V^{\mathbb{C}}$ endlichdimensional und es gilt

$$\dim_{\mathbb{C}}(V^{\mathbb{C}}) = \dim_{\mathbb{R}}(V).$$

IV.2.23. PROPOSITION. *Seien $\phi: V' \xrightarrow{\cong} V$ und $\psi: W \xrightarrow{\cong} W'$ zwei lineare Isomorphismen. Eine lineare Abbildung $\varphi: V \rightarrow W$ hat genau dann endlichen Rang, wenn $\psi \circ \varphi \circ \phi: V' \rightarrow W'$ endlichen Rang hat und in diesem Fall gilt*

$$\text{rank}(\psi \circ \varphi \circ \phi) = \text{rank}(\varphi).$$

BEWEIS. Da ϕ surjektiv ist, gilt $\phi(V') = V$ und somit

$$\text{img}(\psi \circ \varphi \circ \phi) = (\psi \circ \varphi \circ \phi)(V') = \psi(\varphi(\phi(V'))) = \psi(\varphi(V)) = \psi(\text{img}(\varphi)).$$

Wegen der Injektivität von ψ liefert die Einschränkung einen Isomorphismus

$$\psi|_{\text{img}(\varphi)}: \text{img}(\varphi) \xrightarrow{\cong} \psi(\text{img}(\varphi)).$$

Zusammen folgt $\text{img}(\psi \circ \varphi \circ \phi) \cong \text{img}(\varphi)$, also $\text{rank}(\psi \circ \varphi \circ \phi) = \text{rank}(\varphi)$. \square

IV.3. Rang von Matrizen. Sei $A \in M_{m \times n}(\mathbb{K})$ eine Matrix und

$$\psi_A: \mathbb{K}^n \rightarrow \mathbb{K}^m, \quad \psi_A(x) = Ax,$$

die damit assoziierte lineare Abbildung. Der Teilraum

$$L := \ker(\psi_A) \subseteq \mathbb{K}^n \tag{IV.4}$$

ist daher genau der *Lösungsraum* des homogenen Systems $Ax = 0$. Der von den Spalten einer Matrix aufgespannte Teilraum wird ihr *Spaltenraum* genannt. Bezeichnen wir die Spalten mit $A = (a_1 | \dots | a_n)$, so stimmt der Spaltenraum,

$$W := \text{img}(\psi_A) = \langle a_1, \dots, a_n \rangle \subseteq \mathbb{K}^m, \tag{IV.5}$$

also mit dem Teilraum jener $y \in \mathbb{K}^m$ überein, für die das Gleichungssystem $Ax = y$ eine Lösung $x \in \mathbb{K}^n$ hat. Unter dem *Zeilenraum* einer Matrix verstehen wir den von den Zeilen aufgespannten Teilraum. Wir fassen die Zeilenvektoren der Matrix A als Elemente des Dualraums, $(\mathbb{K}^n)^* = L(\mathbb{K}^n, \mathbb{K}) \cong M_{1 \times n}(\mathbb{K})$ auf. Bezeichnen wir die Zeilen von A mit $\alpha_1, \dots, \alpha_m$, dann stimmt der Zeilenraum mit L° überein, siehe Satz III.4.10,

$$L^\circ = \ker(\psi_A)^\circ = \text{img}((\psi_A)^t) = \langle \alpha_1, \dots, \alpha_m \rangle \subseteq (\mathbb{K}^n)^*.$$

Wir können den Zeilenraum daher als den Vektorraum aller linearer Gleichungen verstehen, denen L genügt. Schließlich betrachten wir auch

$$W^\circ = \text{img}(\psi_A)^\circ = \ker((\psi_A)^t) \subseteq (\mathbb{K}^m)^*,$$

den Teilraum aller Gleichungen, denen W genügt.

IV.3.1. DEFINITION (Rang einer Matrix). Unter dem *Rang* einer Matrix $A \in M_{m \times n}(\mathbb{K})$ verstehen wir den Rang der damit assoziierten linearen Abbildung $\psi_A: \mathbb{K}^n \rightarrow \mathbb{K}^m$, $\psi_A(x) = Ax$, d.h.

$$\text{rank}(A) := \text{rank}(\psi_A).$$

Nach Definition gilt daher $\dim(W) = \text{rank}(A)$. Der Rang bestimmt auch die Dimensionen der anderen Teilräume, die wir oben betrachtet haben:

IV.3.2. SATZ (Rang). Sei $A \in M_{m \times n}(\mathbb{K})$ und $k = \text{rank}(A)$. Dann gilt:

- (a) $\dim(L) = n - k$, d.h. der Lösungsraum ist $(n - k)$ -dimensional.
- (b) $\dim(W) = k$, d.h. der Spaltenraum ist k -dimensional. Die Matrix A besitzt daher k linear unabhängige Spalten, und je $k + 1$ Spalten sind linear abhängig.
- (c) $\dim(L^\circ) = k$, d.h. der Zeilenraum ist k -dimensional. Die Matrix A besitzt daher k linear unabhängige Zeilen, und je $k + 1$ Zeilen sind linear abhängig. Jedes minimale lineare Gleichungssystem für L besteht daher aus genau k Gleichungen.
- (d) $\dim(W^\circ) = m - k$. Jedes minimale lineare Gleichungssystem für W besteht daher aus genau $m - k$ Gleichungen.

BEWEIS. Behauptung (b) ist trivial, vgl. Definition IV.3.1 und (IV.5). Behauptung (a) folgt aus Korollar IV.2.9, denn mit (IV.4) erhalten wir

$$\dim(L) = \dim(\ker(\psi_A)) = \dim(\mathbb{K}^n) - \dim(\text{img}(\psi_A)) = n - \text{rank}(A) = n - k.$$

Mit Korollar IV.2.8 und Korollar IV.2.16 folgt nun

$$\dim(L^\circ) = \dim(\mathbb{K}^n/L) = \dim(\mathbb{K}^n) - \dim(L) = n - (n - k) = k.$$

Nach Bemerkung IV.2.18 lässt sich L als Lösungsmenge eines linearen Gleichungssystems mit k Gleichungen beschreiben, und weniger als k lineare Gleichungen reichen nicht aus. D.h. jedes minimale lineare Gleichungssystem für L muss aus genau k Gleichungen bestehen. Damit ist auch (c) gezeigt. Analog erhalten wir $\dim(W^\circ) = \dim(\mathbb{K}^m/W) = \dim(\mathbb{K}^m) - \dim(W) = m - k$, und somit (d). \square

Der Rang wurde als Dimension des Spaltenraums definiert, er wird daher manchmal auch als *Spaltenrang* bezeichnet. Unter dem *Zeilenrang* verstehen wir die Dimension des von den Zeilen aufgespannten Teilraums. Aus dem eben bewiesenen Satz IV.3.2(b)&(c) folgt, dass diese beiden Begriffe übereinstimmen. Wir wollen dies nun nochmals auf direktere Art zeigen:

IV.3.3. SATZ (Rang der Transponierten). Für jede Matrix $A \in M_{m \times n}(\mathbb{K})$ gilt:

$$\text{rank}(A^t) = \text{rank}(A).$$

Der Zeilenrang stimmt daher stets mit dem Spaltenrang überein. Weiters gilt:

$$0 \leq \text{rank}(A) \leq \min(n, m).$$

BEWEIS. Betrachte die lineare Abbildung $\psi_A: \mathbb{K}^n \rightarrow \mathbb{K}^m$, $\psi_A(x) = Ax$, und die duale Abbildung $(\psi_A)^t: (\mathbb{K}^m)^* \rightarrow (\mathbb{K}^n)^*$. In Bemerkung III.4.5 haben wir Isomorphismen $\phi_{\mathbb{K}^m}: \mathbb{K}^m \rightarrow (\mathbb{K}^m)^*$ und $\phi_{\mathbb{K}^n}: \mathbb{K}^n \rightarrow (\mathbb{K}^n)^*$ konstruiert, sodass

$$\psi_{A^t} = \phi_{\mathbb{K}^n}^{-1} \circ (\psi_A)^t \circ \phi_{\mathbb{K}^m}.$$

Mit Proposition IV.2.23 und Satz IV.2.20 erhalten wir daher

$$\begin{aligned} \text{rank}(A^t) &= \text{rank}(\psi_{A^t}) = \text{rank}(\phi_{\mathbb{K}^n}^{-1} \circ (\psi_A)^t \circ \phi_{\mathbb{K}^m}) \\ &= \text{rank}((\psi_A)^t) = \text{rank}(\psi_A) = \text{rank}(A). \end{aligned}$$

Da die Zeilen von A gerade die Spalten von A^t bilden, stimmen Zeilen- und Spaltenrang also überein. Da $\text{img}(\psi_A) \subseteq \mathbb{K}^m$ gilt $\text{rank}(A) = \dim(\text{img}(\psi_A)) \leq \dim(\mathbb{K}^m) = m$ und analog $\text{rank}(A^t) \leq n$. Zusammen mit $\text{rank}(A^t) = \text{rank}(A)$ folgt daraus $\text{rank}(A) \leq \min(n, m)$. \square

IV.3.4. KOROLLAR. Für $A \in M_{m \times n}(\mathbb{K})$ sind folgende Aussagen äquivalent:

- (a) Die Matrix A besitzt eine Rechtsinverse $A' \in M_{n \times m}(\mathbb{K})$, d.h. $AA' = I_m$.
- (b) Die lineare Abbildung $\psi_A: \mathbb{K}^n \rightarrow \mathbb{K}^m$, $\psi_A(x) = Ax$, ist surjektiv.
- (c) Das Gl.system $Ax = y$ hat für jedes $y \in \mathbb{K}^m$ mindestens eine Lösung $x \in \mathbb{K}^n$.
- (d) Die Spalten von A erzeugen \mathbb{K}^m .
- (e) Die Matrix A hat m linear unabhängige Spalten.
- (f) Die Zeilen von A erzeugen einen m -dimensionalen Teilraum.
- (g) Die Zeilen von A sind linear unabhängig.
- (h) Es gilt $\text{rank}(A) = m$.

In diesem Fall muss $m \leq n$ gelten.

BEWEIS. Gilt $AA' = I_m$, so erhalten wir für die assoziierten linearen Abbildungen $\psi_A \circ \psi_{A'} = \psi_{AA'} = \psi_{I_m} = \text{id}_{\mathbb{K}^m}$, also muss ψ_A surjektiv sein. Ist umgekehrt ψ_A surjektiv, dann existiert eine lineare Rechtsinverse $\varphi: \mathbb{K}^m \rightarrow \mathbb{K}^n$, d.h. $\psi_A \circ \varphi = \text{id}_{\mathbb{K}^m}$, siehe Korollar III.3.27(b). Bezeichnet nun A' die entsprechende Matrix, $\varphi = \psi_{A'}$, dann folgt $\psi_{AA'} = \psi_A \circ \psi_{A'} = \psi_A \circ \varphi = \text{id}_{\mathbb{K}^m} = \psi_{I_m}$, also $AA' = I_m$. Damit ist die Äquivalenz (a) \Leftrightarrow (b) gezeigt. Die Äquivalenz (b) \Leftrightarrow (c) ist trivial. Die Äquivalenz (b) \Leftrightarrow (d) haben wir bereits früher fest gehalten, vgl. Bemerkung III.1.18. Die Äquivalenz (d) \Leftrightarrow (h) folgt aus der Definition des Rangs. Die Äquivalenz (d) \Leftrightarrow (f) folgt aus Satz IV.3.3. Die Implikation (d) \Rightarrow (e) folgt aus der Tatsache, dass jedes (endliche) Erzeugendensystem eine Basis enthält, siehe Proposition III.3.17. Die umgekehrte Implikation (e) \Rightarrow (d) folgt aus Korollar IV.1.17. Die Äquivalenz (f) \Leftrightarrow (g) folgt aus Korollar IV.1.17, denn A hat genau m Zeilen. Damit ist die Äquivalenz aller Eigenschaften gezeigt. Aus (e) folgt auch sofort der Zusatz $m \leq n$. \square

IV.3.5. KOROLLAR. Für $A \in M_{m \times n}(\mathbb{K})$ sind folgende Aussagen äquivalent:

- (a) Die Matrix A besitzt eine Linksinverse $A' \in M_{n \times m}(\mathbb{K})$, d.h. $A'A = I_n$.
- (b) Die lineare Abbildung $\psi_A: \mathbb{K}^n \rightarrow \mathbb{K}^m$, $\psi_A(x) = Ax$, ist injektiv.

- (c) Das Gl.system $Ax = y$ hat für jedes $y \in \mathbb{K}^m$ höchstens eine Lösung $x \in \mathbb{K}^n$.
 (d) Die Spalten von A erzeugen einen n -dimensionalen Teilraum.
 (e) Die Spalten von A sind linear unabhängig.
 (f) Die Zeilen von A erzeugen $M_{1 \times n}(\mathbb{K})$.
 (g) Die Matrix A hat n linear unabhängige Zeilen.
 (h) Es gilt $\text{rank}(A) = n$.

In diesem Fall muss $n \leq m$ gelten.

Der Beweis kann analog zu dem des vorangehenden Korollars geführt werden und sei den LeserInnen überlassen.

IV.3.6. BEMERKUNG. Offenbar gilt $\text{rank}(A) = 0$ genau dann, wenn $A = 0$.

Wir widmen uns nun der Berechnung des Rangs einer Matrix.

IV.3.7. DEFINITION (Elementare Zeilen- und Spaltenumformungen). Unter einer *elementaren Zeilenumformung* einer Matrix verstehen wir eine der folgenden Modifikationen:

- (I) Vertauschen zweier Zeilen.
 (II) Multiplikation einer Zeile mit einem Skalar $\lambda \in \mathbb{K}$, $\lambda \neq 0$.
 (III) Addition eines Vielfachen einer Zeile zu einer anderen Zeile.

Unter einer *elementaren Spaltenumformung* einer Matrix verstehen wir eine der folgenden Modifikationen:

- (I) Vertauschen zweier Spalten.
 (II) Multiplikation einer Spalte mit einem Skalar $\lambda \in \mathbb{K}$, $\lambda \neq 0$.
 (III) Addition eines Vielfachen einer Spalte zu einer anderen Spalte.

Jede elementare Spaltenumformung kann durch Multiplikation von rechts mit einer invertierbaren Matrix S beschrieben werden, $A \rightsquigarrow AS$. Bei elementaren Spaltenumformungen vom Typ I ist diese Matrix von der Form

$$S = S_I^{i;\lambda} = \begin{pmatrix} \ddots & & & & \\ & \ddots & & & \\ & & 1 & & \\ & & & \lambda & \\ & & & & \ddots \end{pmatrix} = I + (\lambda - 1)E_{i,i},$$

wobei $\lambda \neq 0$, und i bezeichnet die Nummer jener Spalte, die mit λ multipliziert wird. Beachte, dass diese Matrizen invertierbar sind, $(S_I^{i;\lambda})^{-1} = S_I^{i;1/\lambda}$. Bei elementaren Spaltenumformungen vom Typ II hat S die Gestalt

$$S = S_{II}^{i,j;\mu} = \begin{pmatrix} \ddots & & & & \\ & \ddots & & & \\ & & 1 & & \\ & & & \ddots & \\ & & & & 1 \end{pmatrix} = I + \mu E_{i,j},$$

wobei $\mu \in \mathbb{K}$, $i \neq j$, und i bezeichnet die Nummer jener Spalte, die μ -mal zur j -ten Spalte addiert wird. Auch diese Matrizen sind invertierbar, $(S_{II}^{i,j;\mu})^{-1} = S_{II}^{i,j;-\mu}$.

$\tilde{A} = T_N \cdots T_2 T_1 A$, also $\tilde{A} = TA$, wobei $T = T_N \cdots T_2 T_1 \in \text{GL}_m(\mathbb{K})$. Wir erhalten aber auch $\tilde{A}^t = A^t T_1^t T_2^t \cdots T_N^t$, d.h. \tilde{A}^t entsteht aus A^t durch eine Folge elementarer Spaltenumformungen, siehe (IV.6). Nach Lemma IV.3.8 gilt daher $\text{rank}(\tilde{A}^t) = \text{rank}(A^t)$ und der Spaltenraum von \tilde{A}^t stimmt mit dem Spaltenraum von A^t überein. Mit Satz IV.3.3 folgt $\text{rank}(\tilde{A}) = \text{rank}(A)$. Da der Spaltenraum der Transponierten gerade der Zeilenraum der ursprünglichen Matrix ist, müssen auch die Zeilenräume von \tilde{A} und A gleich sein. Aus der Invertierbarkeit von T folgt sofort $\tilde{A}x = 0 \Leftrightarrow Ax = 0$, $x \in \mathbb{K}^n$. \square

IV.3.10. DEFINITION (Zeilen- und Spaltenstufenform). Wir sagen eine Matrix $\tilde{A} \in M_{m \times n}(\mathbb{K})$ hat *Zeilenstufenform*, wenn sie die Gestalt

$$\begin{pmatrix} 0 \cdots 0 & \tilde{A}_{1j_1} & * \cdots * & * & * \cdots * & * & \cdots & * \cdots * & * & * \cdots * \\ 0 \cdots 0 & 0 & 0 \cdots 0 & \tilde{A}_{2j_2} & * \cdots * & * & \cdots & * \cdots * & * & * \cdots * \\ 0 \cdots 0 & 0 & 0 \cdots 0 & 0 & 0 \cdots 0 & \tilde{A}_{3j_3} & \cdots & * \cdots * & * & * \cdots * \\ 0 \cdots 0 & 0 & 0 \cdots 0 & 0 & 0 \cdots 0 & 0 & \cdots & * \cdots * & * & * \cdots * \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 \cdots 0 & 0 & 0 \cdots 0 & 0 & 0 \cdots 0 & 0 & \cdots & * \cdots * & * & * \cdots * \\ 0 \cdots 0 & 0 & 0 \cdots 0 & 0 & 0 \cdots 0 & 0 & \cdots & 0 \cdots 0 & \tilde{A}_{kj_k} & * \cdots * \\ 0 \cdots 0 & 0 & 0 \cdots 0 & 0 & 0 \cdots 0 & 0 & \cdots & 0 \cdots 0 & 0 & 0 \cdots 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 \cdots 0 & 0 & 0 \cdots 0 & 0 & 0 \cdots 0 & 0 & \cdots & 0 \cdots 0 & 0 & 0 \cdots 0 \end{pmatrix}$$

hat, wobei $1 \leq j_1 < j_2 < \cdots < j_k \leq n$, und die Einträge $\tilde{A}_{1j_1}, \dots, \tilde{A}_{kj_k}$ alle verschieden von Null sind. Die Matrix hat *reduzierte Zeilenstufenform*, falls sie folgende Gestalt hat:

$$\begin{pmatrix} 0 \cdots 0 & 1 & * \cdots * & 0 & * \cdots * & 0 & \cdots & * \cdots * & 0 & * \cdots * \\ 0 \cdots 0 & 0 & 0 \cdots 0 & 1 & * \cdots * & 0 & \cdots & * \cdots * & 0 & * \cdots * \\ 0 \cdots 0 & 0 & 0 \cdots 0 & 0 & 0 \cdots 0 & 1 & \cdots & * \cdots * & 0 & * \cdots * \\ 0 \cdots 0 & 0 & 0 \cdots 0 & 0 & 0 \cdots 0 & 0 & \cdots & * \cdots * & 0 & * \cdots * \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 \cdots 0 & 0 & 0 \cdots 0 & 0 & 0 \cdots 0 & 0 & \cdots & * \cdots * & 0 & * \cdots * \\ 0 \cdots 0 & 0 & 0 \cdots 0 & 0 & 0 \cdots 0 & 0 & \cdots & 0 \cdots 0 & 1 & * \cdots * \\ 0 \cdots 0 & 0 & 0 \cdots 0 & 0 & 0 \cdots 0 & 0 & \cdots & 0 \cdots 0 & 0 & 0 \cdots 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 \cdots 0 & 0 & 0 \cdots 0 & 0 & 0 \cdots 0 & 0 & \cdots & 0 \cdots 0 & 0 & 0 \cdots 0 \end{pmatrix}$$

Wir sagen \tilde{A} hat (*reduzierte*) *Spaltenstufenform*, wenn \tilde{A}^t (*reduzierte*) Zeilenstufenform hat. Eine Matrix \tilde{A} ist also genau dann von Spaltenstufenform, wenn sie folgende Gestalt hat

$$\begin{pmatrix} 0 & 0 & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \cdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & 0 & 0 \\ \tilde{A}_{i_1 1} & 0 & \cdots & 0 & 0 & \cdots 0 \\ * & 0 & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \cdots & \vdots & \vdots & \vdots \\ * & 0 & \cdots & 0 & 0 & 0 \\ * & \tilde{A}_{i_2 2} & \cdots & 0 & 0 & \cdots 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ * & * & \cdots & \tilde{A}_{i_k k} & 0 & \cdots 0 \\ * & * & \cdots & * & 0 & 0 \\ \vdots & \vdots & \cdots & \vdots & \vdots & \vdots \\ * & * & \cdots & * & 0 & 0 \end{pmatrix}$$

wobei $1 \leq i_1 < i_2 < \dots < i_k \leq m$ und die Eintragungen $\tilde{A}_{i_1 1}, \dots, \tilde{A}_{i_k k}$ alle verschieden von Null sind.

IV.3.11. SATZ (Elimination). *Jede Matrix $A \in M_{m \times n}(\mathbb{K})$ kann durch endlich viele elementare Zeilenumformungen auf (reduzierte) Zeilenstufenform \tilde{A} gebracht werden. Sind dabei $1 \leq j_1 < j_2 < \dots < j_k \leq n$ wie in Definition IV.3.10 dann gilt:*

- (a) $\text{rank}(A) = k$.
- (b) Die ersten k Zeilen der Zeilenstufenform \tilde{A} bilden eine Basis des Zeilenraums von A , und liefern daher ein minimales Gleichungssystem für den Lösungsraum $L = \{x \in \mathbb{K}^n \mid Ax = 0\}$.
- (c) Die Einheitsvektoren e_{j_1}, \dots, e_{j_k} bilden eine Basis für einen zu L komplementären Teilraum L' . Die Zeilenvektoren e_j^t , $j \in \{1, \dots, n\} \setminus \{j_1, \dots, j_k\}$, liefern ein minimales Gleichungssystem für L' .
- (d) Die Spalten von A mit den Nummern j_1, \dots, j_k bilden eine Basis des Spaltenraums von A .
- (e) Ist die Zeilenstufenform \tilde{A} reduziert, so bilden die Vektoren

$$e_j - \sum_{\{l \mid 1 \leq j_l < j\}} \tilde{A}_{lj} e_{j_l}, \quad j \in \{1, \dots, n\} \setminus \{j_1, \dots, j_k\},$$

eine Basis des Lösungsraums L .

BEWEIS. Wir bringen die Matrix A spaltenweise von links nach rechts auf reduzierte Zeilenstufenform. Induktiv nehmen wir an die ersten $q - 1$ Spalten von A sind schon in reduzierter Zeilenstufenform. Sei p minimal für die Eigenschaft $A_{p1} = \dots = A_{p,q-1} = 0$. Sind die Elemente A_{pq}, \dots, A_{mq} alle Null, dann sind auch die ersten q Spalten von A in reduzierter Zeilenstufenform, und der Induktionsschritt erledigt. O.B.d.A. sei also einer der Einträge A_{pq}, \dots, A_{mq} verschieden von Null. Durch Vertauschen zweier Zeilen können wir weiters $A_{pq} \neq 0$ erreichen. Durch Multiplikation der p -ten Zeile mit $1/A_{pq}$, dürfen wir weiters annehmen, dass $A_{pq} = 1$. Durch Addition geeigneter Vielfacher der p -ten Zeile zu den restlichen Zeilen können wir schließlich auch $A_{iq} = 0$, für alle $i \neq p$ erreichen. Beachte, dass die verwendeten Zeilenumformungen die ersten $q - 1$ Spalten von A unverändert lassen. Wir erhalten also eine Matrix, deren ersten q Spalten in reduzierter Zeilenstufenform vorliegen. Damit ist der Induktionsschritt gezeigt, also lässt sich A durch elementare Zeilenumformungen auf reduzierte Zeilenstufenform \tilde{A} bringen. Nach Lemma IV.3.9 gilt $\text{rank}(\tilde{A}) = \text{rank}(A)$, die Zeilenräume von A und \tilde{A} sind gleich und auch die Lösungsräume von A und \tilde{A} stimmen überein.

Im Fall $\tilde{A} = A$ lassen sich die Aussagen (a), (b), (c) und (e) aufgrund der einfachen Zeilenstufenform sofort ablesen. Dies sind aber alles nur Aussagen über den Zeilen- bzw. Lösungsraum. Da auch im allgemeinen Fall Zeilen- und Lösungsraum der beiden Matrizen A und \tilde{A} übereinstimmen, bleiben sie für A richtig.

Nun zur verbleibenden Behauptung (d): Der Spaltenraum von A hat Dimension k , siehe (a) und Satz IV.3.2(b). Es genügt daher zu zeigen, dass die Spalten mit Nummern j_1, \dots, j_k linear unabhängig sind, siehe Korollar IV.1.17. Es bezeichne B die Matrix die wir aus A erhalten indem wir alle Spalten bis auf die mit Nummern j_1, \dots, j_k streichen. Es genügt $\text{rank}(B) = k$ zu zeigen, siehe Satz IV.3.5. Wenden wir die selben Zeilenumformungen, die uns von A zu \tilde{A} geführt haben, auf B an, so erhalten wir eine Matrix \tilde{B} , deren Spalten gerade die Spalten von \tilde{A} mit Nummern j_1, \dots, j_k sind. Nach Lemma IV.3.9 genügt es $\text{rank}(\tilde{B}) = k$ zu zeigen. Wegen der einfachen Gestalt von \tilde{A} lässt sich dies aber sofort ablesen. \square

IV.3.12. SATZ (Elimination). *Jede Matrix $A \in M_{m \times n}(\mathbb{K})$ kann durch endlich viele elementare Spaltenumformungen auf (reduzierte) Spaltenstufenform \tilde{A} gebracht werden. Sind dabei $1 \leq i_1 < i_2 < \dots < i_k \leq m$ wie in Definition IV.3.10, dann gilt:*

- (a) $\text{rank}(A) = k$.
- (b) Die ersten k Spalten der Spaltenstufenform \tilde{A} bilden eine Basis des Spaltenraums W von A .
- (c) Die Spaltenvektoren e_i , $i \in \{1, \dots, m\} \setminus \{i_1, \dots, i_k\}$, bilden eine Basis für einen zu W komplementären Teilraum W' . Die Zeilenvektoren $e_{i_1}^t, \dots, e_{i_k}^t$ liefern ein minimales Gleichungssystem für W' .
- (d) Die Zeilen von A mit den Nummern i_1, \dots, i_k bilden eine Basis des Zeilenraums von A und liefern daher ein minimales Gleichungssystem für den Lösungsraum $L = \{x \in \mathbb{K}^n : Ax = 0\}$.
- (e) Ist die Spaltenstufenform \tilde{A} reduziert, so bilden die Zeilenvektoren

$$e_i^t - \sum_{\{l \mid 1 \leq i_l < i\}} \tilde{A}_{il} e_{i_l}^t, \quad i \in \{1, \dots, m\} \setminus \{i_1, \dots, i_k\},$$

eine Basis von W° und liefern daher ein minimales Gleichungssystem für den Spaltenraum von A .

BEWEIS. Wir führen die erste Aussage auf den vorangehenden Satz IV.3.11 zurück. Dieser besagt, dass A^t durch elementare Zeilenumformungen auf reduzierte Zeilenstufenform gebracht werden kann. Es existieren daher invertierbare Matrizen $T_1, \dots, T_N \in \text{GL}_n(\mathbb{K})$ wie in (IV.6), sodass $T_N \cdots T_1 A^t$ reduzierte Zeilenstufenform hat. Also ist $AT_1^t \cdots T_N^t$ eine Matrix in reduzierter Spaltenstufenform. Da die Multiplikation von rechts mit T_i^t gerade einer elementaren Spaltenumformung entspricht sehen wir, dass A durch N elementare Spaltenumformungen auf reduzierte Spaltenstufenform \tilde{A} gebracht werden kann. Nach Lemma IV.3.8 gilt $\text{rank}(\tilde{A}) = \text{rank}(A)$, und auch der Spaltenraum von \tilde{A} stimmt mit dem Spaltenraum von A überein.

Wie im Beweis von Satz IV.3.11 lassen sich die Aussagen (a), (b), (c) und (e) für die (reduzierte) Spaltenstufenform \tilde{A} sofort ablesen. Da es sich dabei um

Aussagen über den Spaltenraum handelt, bleiben sie für A richtig, denn die Spaltenräume von A und \tilde{A} sind gleich. Die verbleibende Behauptung (d) lässt sich wie zuvor zeigen, oder kann durch Übergang zur Transponierten aus Satz IV.3.11 abgeleitet werden. \square

IV.3.13. BEISPIEL. Wir wollen zunächst den Rang der reellen Matrix

$$A = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 4 & 5 \\ 0 & 0 & 2 & 3 & 4 & 1 & 3 & 4 \\ 0 & 0 & 2 & 3 & 4 & 4 & 7 & 5 \\ 0 & 0 & 4 & 6 & 8 & 2 & 14 & 18 \\ 0 & 0 & 6 & 9 & 12 & 9 & 17 & 14 \\ 0 & 0 & -2 & -3 & -4 & -4 & -3 & 0 \end{pmatrix} \in M_{6 \times 8}(\mathbb{R})$$

bestimmen. Durch Vertauschen der ersten und dritten Zeile erhalten wir:

$$\begin{pmatrix} 0 & 0 & 2 & 3 & 4 & 1 & 3 & 4 \\ 0 & 0 & 0 & 0 & 0 & 0 & 4 & 5 \\ 0 & 0 & 2 & 3 & 4 & 4 & 7 & 5 \\ 0 & 0 & 4 & 6 & 8 & 2 & 14 & 18 \\ 0 & 0 & 6 & 9 & 12 & 9 & 17 & 14 \\ 0 & 0 & -2 & -3 & -4 & -4 & -3 & 0 \end{pmatrix}$$

Durch Addition geeigneter Vielfacher der ersten Zeile zu den anderen Zeilen:

$$\begin{pmatrix} 0 & 0 & 2 & 3 & 4 & 1 & 3 & 4 \\ 0 & 0 & 0 & 0 & 0 & 0 & 4 & 5 \\ 0 & 0 & 0 & 0 & 0 & 3 & 4 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 8 & 10 \\ 0 & 0 & 0 & 0 & 0 & 6 & 8 & 2 \\ 0 & 0 & 0 & 0 & 0 & -3 & 0 & 4 \end{pmatrix}$$

Vertauschen der zweiten und dritten Zeile liefert:

$$\begin{pmatrix} 0 & 0 & 2 & 3 & 4 & 1 & 3 & 4 \\ 0 & 0 & 0 & 0 & 0 & 3 & 4 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 4 & 5 \\ 0 & 0 & 0 & 0 & 0 & 0 & 8 & 10 \\ 0 & 0 & 0 & 0 & 0 & 6 & 8 & 2 \\ 0 & 0 & 0 & 0 & 0 & -3 & 0 & 4 \end{pmatrix}$$

Durch Addition geeigneter Vielfacher der zweiten Zeile zu den letzten beiden Zeilen erhalten wir:

$$\begin{pmatrix} 0 & 0 & 2 & 3 & 4 & 1 & 3 & 4 \\ 0 & 0 & 0 & 0 & 0 & 3 & 4 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 4 & 5 \\ 0 & 0 & 0 & 0 & 0 & 0 & 8 & 10 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 4 & 5 \end{pmatrix}$$

Durch Addition geeigneter Vielfacher der dritten Zeile zur vierten und sechsten Zeile erhalten wir schließlich eine Matrix in Zeilenstufenform:

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

Nach Satz IV.3.11(a) gilt daher $\text{rank}(A) = 3$. Aus dieser Rechnung lässt sich noch viel mehr über A sagen. Nach Satz IV.3.11(b) bildet

$$\begin{aligned} 2x_3 + 3x_4 + 4x_5 + x_6 + 3x_7 + 4x_8 &= 0 \\ 3x_6 + 4x_7 + x_8 &= 0 \\ 4x_7 + 5x_8 &= 0 \end{aligned} \quad (\text{IV.7})$$

ein minimales Gleichungssystem für den Lösungsraum $L = \{x \in \mathbb{R}^8 \mid Ax = 0\}$, und es gilt $\dim(L) = 5$, siehe Satz IV.3.2(a). Nach Satz IV.3.11(c) ist $L' := \langle e_3, e_6, e_7 \rangle$ ein zu L komplementärer Teilraum, $\dim(L') = 3$, und dieser lässt sich durch das minimale Gleichungssystem $x_1 = x_2 = x_4 = x_5 = x_8 = 0$ beschreiben. Nach Satz IV.3.11(d) bilden die folgenden Spalten von A

$$a_3 = \begin{pmatrix} 0 \\ 2 \\ 2 \\ 4 \\ 6 \\ -2 \end{pmatrix}, \quad a_6 = \begin{pmatrix} 0 \\ 1 \\ 4 \\ 2 \\ 9 \\ -4 \end{pmatrix}, \quad a_7 = \begin{pmatrix} 4 \\ 3 \\ 7 \\ 14 \\ 17 \\ -3 \end{pmatrix},$$

eine Basis des Spaltenraums von A . Bringen wir A durch weitere Zeilenumformungen auf reduzierte Zeilenstufenform,

$$\begin{aligned} \begin{pmatrix} 0 & 0 & 2 & 3 & 4 & 1 & 3 & 4 \\ 0 & 0 & 0 & 0 & 0 & 3 & 4 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 4 & 5 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} &\rightsquigarrow \begin{pmatrix} 0 & 0 & 1 & 3/2 & 2 & 1/2 & 3/2 & 2 \\ 0 & 0 & 0 & 0 & 0 & 1 & 4/3 & 1/3 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 5/4 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \\ &\rightsquigarrow \begin{pmatrix} 0 & 0 & 1 & 3/2 & 2 & 1/2 & 0 & 1/8 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & -4/3 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 5/4 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} &\rightsquigarrow \begin{pmatrix} 0 & 0 & 1 & 3/2 & 2 & 0 & 0 & 19/24 \\ 0 & 0 & 0 & 0 & 1 & 0 & -4/3 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 5/4 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \end{aligned}$$

dann folgt aus Satz IV.3.11(e), dass die Vektoren

$$b_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \quad b_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \quad b_3 = \begin{pmatrix} 0 \\ 0 \\ -3/2 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \quad b_4 = \begin{pmatrix} 0 \\ 0 \\ -2 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \quad b_5 = \begin{pmatrix} 0 \\ 0 \\ -19/24 \\ 0 \\ 4/3 \\ -5/4 \\ 1 \end{pmatrix}$$

eine Basis des Lösungsraums $L = \{x \in \mathbb{R}^8 \mid Ax = 0\}$ bilden. Somit ist $\phi: \mathbb{R}^5 \xrightarrow{\cong} L$,

$$\phi \begin{pmatrix} s_1 \\ s_2 \\ s_3 \\ s_4 \\ s_5 \end{pmatrix} = s_1 \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} + s_2 \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} + s_3 \begin{pmatrix} 0 \\ 0 \\ -3/2 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} + s_4 \begin{pmatrix} 0 \\ 0 \\ -2 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} + s_5 \begin{pmatrix} 0 \\ 0 \\ -19/24 \\ 0 \\ 0 \\ 4/3 \\ -5/4 \\ 1 \end{pmatrix}$$

ein Isomorphismus, d.h. eine Parameterdarstellung von L . Schließlich erhalten wir aus der reduzierten Zeilenstufenform folgendes etwas einfachere minimale Gleichungssystem für L , vgl. (IV.7):

$$\begin{aligned} x_3 + \frac{3}{2}x_4 + 2x_5 + \frac{19}{24}x_8 &= 0 \\ x_6 - \frac{4}{3}x_8 &= 0 \\ x_7 + \frac{5}{4}x_8 &= 0 \end{aligned}$$

IV.3.14. BEISPIEL. Wir wollen den Rang der reellen Matrix

$$A = \begin{pmatrix} 0 & 0 & 3 & 7 \\ 1 & 5 & 8 & 10 \\ 2 & 12 & 22 & 29 \\ 3 & 13 & 21 & 32 \end{pmatrix} \in M_{4 \times 4}(\mathbb{R})$$

bestimmen. Durch entsprechende Zeilenumformungen erhalten wir:

$$\begin{aligned} \begin{pmatrix} 0 & 0 & 3 & 7 \\ 1 & 5 & 8 & 10 \\ 2 & 12 & 22 & 29 \\ 3 & 13 & 21 & 32 \end{pmatrix} &\rightsquigarrow \begin{pmatrix} 1 & 5 & 8 & 10 \\ 0 & 0 & 3 & 7 \\ 2 & 12 & 22 & 29 \\ 3 & 13 & 21 & 32 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1 & 5 & 8 & 10 \\ 0 & 0 & 3 & 7 \\ 0 & 2 & 6 & 9 \\ 0 & -2 & -3 & 2 \end{pmatrix} \\ &\rightsquigarrow \begin{pmatrix} 1 & 5 & 8 & 10 \\ 0 & 2 & 6 & 9 \\ 0 & 0 & 3 & 7 \\ 0 & -2 & -3 & 2 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1 & 5 & 8 & 10 \\ 0 & 2 & 6 & 9 \\ 0 & 0 & 3 & 7 \\ 0 & 0 & 3 & 11 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1 & 5 & 8 & 10 \\ 0 & 2 & 6 & 9 \\ 0 & 0 & 3 & 7 \\ 0 & 0 & 0 & 4 \end{pmatrix} \end{aligned}$$

Somit $\text{rank}(A) = 4$, die Matrix A ist daher invertierbar, siehe Korollar IV.3.5.

IV.3.15. BEISPIEL. Wir wollen den Rang der komplexen Matrix

$$A = \begin{pmatrix} \mathbf{i} & 7 & 1 - \mathbf{i} \\ 2\mathbf{i} & 14 & 3 - \mathbf{i} \\ -1 & 2 + 4\mathbf{i} & 1 + 4\mathbf{i} \\ 1 - \mathbf{i} & -4 - 5\mathbf{i} & 7\mathbf{i} \end{pmatrix} \in M_{4 \times 3}(\mathbb{C})$$

bestimmen. Durch Zeilenumformungen erhalten wir:

$$\begin{aligned} \begin{pmatrix} \mathbf{i} & 7 & 1 - \mathbf{i} \\ 2\mathbf{i} & 14 & 3 - \mathbf{i} \\ -1 & 2 + 4\mathbf{i} & 1 + 4\mathbf{i} \\ 1 - \mathbf{i} & -4 - 5\mathbf{i} & 7\mathbf{i} \end{pmatrix} &\rightsquigarrow \begin{pmatrix} \mathbf{i} & 7 & 1 - \mathbf{i} \\ 0 & 0 & 1 + \mathbf{i} \\ 0 & 2 - 3\mathbf{i} & 3\mathbf{i} \\ 0 & 3 + 2\mathbf{i} & 2 + 7\mathbf{i} \end{pmatrix} \rightsquigarrow \begin{pmatrix} \mathbf{i} & 7 & 1 - \mathbf{i} \\ 0 & 2 - 3\mathbf{i} & 3\mathbf{i} \\ 0 & 0 & 1 + \mathbf{i} \\ 0 & 3 + 2\mathbf{i} & 2 + 7\mathbf{i} \end{pmatrix} \\ &\rightsquigarrow \begin{pmatrix} \mathbf{i} & 7 & 1 - \mathbf{i} \\ 0 & 2 - 3\mathbf{i} & 3\mathbf{i} \\ 0 & 0 & 1 + \mathbf{i} \\ 0 & 0 & 5 + 7\mathbf{i} \end{pmatrix} \rightsquigarrow \begin{pmatrix} \mathbf{i} & 7 & 1 - \mathbf{i} \\ 0 & 2 - 3\mathbf{i} & 3\mathbf{i} \\ 0 & 0 & 1 + \mathbf{i} \\ 0 & 0 & 0 \end{pmatrix} \end{aligned}$$

Es gilt daher $\text{rank}(A) = 3$. Die lineare Abbildung $\psi_A: \mathbb{C}^3 \rightarrow \mathbb{C}^4$, $\psi_A(x) = Ax$, ist daher injektiv. Das Gleichungssystem $Ax = y$ hat daher für jedes $y \in \mathbb{C}^4$ höchstens eine Lösung $x \in \mathbb{C}^3$, siehe Korollar IV.3.5.

IV.3.16. BEISPIEL. Wir wollen den Rang folgender Matrix über dem Körper $\mathbb{K} = \mathbb{Z}_2$ berechnen:

$$A = \begin{pmatrix} 0 & 1 & 0 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 1 & 0 & 1 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 1 & 0 & 1 & 1 & 1 & 1 \end{pmatrix} \in M_{5 \times 6}(\mathbb{Z}_2)$$

Durch Zeilenumformungen erhalten wir:

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

Es gilt daher $\text{rank}(A) = 4$.

IV.3.17. BEISPIEL. Betrachte den von den Vektoren

$$v_1 = \begin{pmatrix} 1 \\ 2 \\ 4 \\ 6 \\ 1 \end{pmatrix}, \quad v_2 = \begin{pmatrix} 2 \\ 4 \\ 8 \\ 15 \\ 6 \end{pmatrix}, \quad v_3 = \begin{pmatrix} 1 \\ 2 \\ 4 \\ 9 \\ 8 \end{pmatrix}, \quad v_4 = \begin{pmatrix} 4 \\ 8 \\ 16 \\ 30 \\ 15 \end{pmatrix}, \quad v_5 = \begin{pmatrix} 3 \\ 6 \\ 12 \\ 24 \\ 12 \end{pmatrix}$$

aufgespannten Teilraum $W := \langle v_1, v_2, v_3, v_4, v_5 \rangle \subseteq \mathbb{R}^5$. Wir wollen nun:

- (1) $\dim(W)$ und eine Basis von W bestimmen.
- (2) Ein minimales Gleichungssystem für W angeben.
- (3) Einen zu W komplementären Teilraum W' bestimmen und diesen durch eine Basis und ein minimales Gleichungssystem beschreiben.
- (4) Eine Basis von W bestimmen, die aus gewissen der Vektoren v_i besteht.

Fassen wir die Vektoren zu einer Matrix zusammen,

$$A = \begin{pmatrix} 1 & 2 & 1 & 4 & 3 \\ 2 & 4 & 2 & 8 & 6 \\ 4 & 8 & 4 & 16 & 12 \\ 6 & 15 & 9 & 30 & 24 \\ 1 & 6 & 8 & 15 & 12 \end{pmatrix},$$

dann ist W der Spaltenraum von A . Durch Spaltenumformungen erhalten wir

$$\begin{pmatrix} 1 & 2 & 1 & 4 & 3 \\ 2 & 4 & 2 & 8 & 6 \\ 4 & 8 & 4 & 16 & 12 \\ 6 & 15 & 9 & 30 & 24 \\ 1 & 6 & 8 & 15 & 12 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 2 & 0 & 0 & 0 & 0 \\ 4 & 0 & 0 & 0 & 0 \\ 6 & 3 & 3 & 6 & 6 \\ 1 & 4 & 7 & 11 & 9 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 2 & 0 & 0 & 0 & 0 \\ 4 & 0 & 0 & 0 & 0 \\ 6 & 3 & 3 & 6 & 6 \\ 1 & 4 & 3 & 3 & 1 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 2 & 0 & 0 & 0 & 0 \\ 4 & 0 & 0 & 0 & 0 \\ 6 & 3 & 3 & 6 & 6 \\ 1 & 4 & 3 & 0 & 0 \end{pmatrix}$$

Nach Satz IV.3.12(b) bilden daher die Vektoren

$$b_1 = \begin{pmatrix} 1 \\ 2 \\ 4 \\ 6 \\ 1 \end{pmatrix}, \quad b_2 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 3 \\ 4 \end{pmatrix}, \quad b_3 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 3 \end{pmatrix}$$

eine Basis von W und es gilt $\dim(W) = 3$. Durch weitere Spaltenumformungen kann die Matrix auf reduzierte Spaltenstufenform gebracht werden:

$$\begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 2 & 0 & 0 & 0 & 0 \\ 4 & 0 & 0 & 0 & 0 \\ 6 & 3 & 0 & 0 & 0 \\ 1 & 4 & 3 & 0 & 0 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 2 & 0 & 0 & 0 & 0 \\ 4 & 0 & 0 & 0 & 0 \\ 6 & 3 & 0 & 0 & 0 \\ 1 & 4 & 1 & 0 & 0 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 2 & 0 & 0 & 0 & 0 \\ 4 & 0 & 0 & 0 & 0 \\ 6 & 3 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 2 & 0 & 0 & 0 & 0 \\ 4 & 0 & 0 & 0 & 0 \\ 6 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 2 & 0 & 0 & 0 & 0 \\ 4 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \end{pmatrix}$$

Wir erhalten so eine Basis von W ,

$$b'_1 = \begin{pmatrix} 1 \\ 2 \\ 4 \\ 0 \\ 0 \end{pmatrix}, \quad b'_2 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \quad b'_3 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{pmatrix},$$

die sehr viele verschwindende Komponenten hat. Insbesondere ist

$$\phi: \mathbb{R}^3 \xrightarrow{\cong} W, \quad \phi \begin{pmatrix} s \\ t \\ u \end{pmatrix} = s \begin{pmatrix} 1 \\ 2 \\ 4 \\ 0 \\ 0 \end{pmatrix} + t \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} + u \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} s \\ 2s \\ 4s \\ t \\ u \end{pmatrix},$$

ein Isomorphismus, d.h. Parameterdarstellung von W . Nach Satz IV.3.12(e) ist

$$\begin{aligned} -2x_1 + x_2 &= 0 \\ -4x_1 + x_3 &= 0 \end{aligned}$$

ein minimales Gleichungssystem für W . Nach Satz IV.3.12(c) bilden die Vektoren

$$e_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \quad e_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 0 \end{pmatrix},$$

eine Basis für einen zu W komplementären Teilraum $W' = \langle e_2, e_3 \rangle$. Ein minimales Gleichungssystem für W' ist:

$$\begin{aligned} x_1 &= 0 \\ x_4 &= 0 \\ x_5 &= 0 \end{aligned}$$

Um die letzte Frage zu beantworten bringen wir A auf Zeilenstufenform:

$$\begin{pmatrix} 1 & 2 & 1 & 4 & 3 \\ 2 & 4 & 2 & 8 & 6 \\ 4 & 8 & 4 & 16 & 12 \\ 6 & 15 & 9 & 30 & 24 \\ 1 & 6 & 8 & 15 & 12 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1 & 2 & 1 & 4 & 3 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 3 & 3 & 6 & 6 \\ 0 & 4 & 7 & 11 & 9 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1 & 2 & 1 & 4 & 3 \\ 0 & 3 & 3 & 6 & 6 \\ 0 & 4 & 7 & 11 & 9 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1 & 2 & 1 & 4 & 3 \\ 0 & 1 & 1 & 2 & 2 \\ 0 & 4 & 7 & 11 & 9 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1 & 2 & 1 & 4 & 3 \\ 0 & 1 & 1 & 2 & 2 \\ 0 & 0 & 3 & 3 & 1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Nach Satz IV.3.11(d) bilden auch die Vektoren v_1, v_2, v_3 eine Basis von W .

IV.3.18. BEISPIEL. Wir wollen verifizieren, dass die Vektoren

$$v_1 = \begin{pmatrix} 7 \\ -1 \\ 2 \\ 0 \\ 3 \\ 1 \end{pmatrix}, \quad v_2 = \begin{pmatrix} 14 \\ -1 \\ 4 \\ 1 \\ 10 \\ 3 \end{pmatrix}, \quad v_3 = \begin{pmatrix} -21 \\ 0 \\ -6 \\ -3 \\ -23 \\ 2 \end{pmatrix},$$

linear unabhängig in \mathbb{R}^6 sind und diese zu einer Basis von \mathbb{R}^6 ergänzen. Mit Hilfe von Spaltenumformungen erhalten wir

$$\begin{pmatrix} 7 & 14 & -21 \\ -1 & -1 & 0 \\ 2 & 4 & -6 \\ 0 & 1 & -3 \\ 3 & 10 & -23 \\ 1 & 3 & 2 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 7 & 0 & 0 \\ -1 & 1 & -3 \\ 2 & 0 & 0 \\ 0 & 1 & -3 \\ 3 & 4 & -14 \\ 1 & 1 & 5 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 7 & 0 & 0 \\ -1 & 1 & 0 \\ 2 & 0 & 0 \\ 0 & 1 & 0 \\ 3 & 4 & 1 \\ 1 & 1 & 8 \end{pmatrix}$$

Die Vektoren v_1, v_2, v_3 sind also linear unabhängig, siehe Korollar IV.3.5. Zusammen mit den Einheitsvektoren e_3, e_4 und e_6 bilden sie eine Basis $v_1, v_2, v_3, e_3, e_4, e_6$ von \mathbb{R}^6 , siehe Satz IV.3.12(c).

IV.3.19. BEISPIEL. Es bezeichne $L \subseteq \mathbb{R}^5$ den Lösungsraum des Systems:

$$\begin{aligned} x_1 + 2x_2 + x_3 + 4x_4 + 11x_5 &= 0 \\ 2x_1 + 4x_2 + 4x_3 + 14x_4 + 30x_5 &= 0 \\ 3x_1 + 6x_2 + 3x_3 + 12x_4 + 34x_5 &= 0 \\ 2x_1 + 4x_2 + 6x_3 + 20x_4 + 41x_5 &= 0 \\ x_1 + 2x_2 - x_3 - 2x_4 + 7x_5 &= 0 \end{aligned} \tag{IV.8}$$

Wir wollen nun:

- (1) $\dim(L)$ und eine Basis von L bestimmen.
- (2) Ein minimales Gleichungssystem für L angeben.
- (3) Einen zu L komplementären Teilraum L' bestimmen und diesen durch eine Basis und ein minimales Gleichungssystem beschreiben.
- (4) Ein minimales Gleichungssystem für L angeben, das aus gewissen der ursprünglichen Gleichungen besteht.

Wir fassen die Koeffizienten des Gleichungssystems zu einer Matrix zusammen,

$$A = \begin{pmatrix} 1 & 2 & 1 & 4 & 11 \\ 2 & 4 & 4 & 14 & 30 \\ 3 & 6 & 3 & 12 & 34 \\ 2 & 4 & 6 & 20 & 41 \\ 1 & 2 & -1 & -2 & 7 \end{pmatrix}.$$

Durch Zeilenumformungen bringen wir A auf reduzierte Zeilenstufenform:

$$\begin{aligned} \begin{pmatrix} 1 & 2 & 1 & 4 & 11 \\ 2 & 4 & 4 & 14 & 30 \\ 3 & 6 & 3 & 12 & 34 \\ 2 & 4 & 6 & 20 & 41 \\ 1 & 2 & -1 & -2 & 7 \end{pmatrix} &\rightsquigarrow \begin{pmatrix} 1 & 2 & 1 & 4 & 11 \\ 0 & 0 & 2 & 6 & 8 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 4 & 12 & 19 \\ 0 & 0 & -2 & -6 & -4 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1 & 2 & 1 & 4 & 11 \\ 0 & 0 & 2 & 6 & 8 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 & 4 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1 & 2 & 1 & 4 & 11 \\ 0 & 0 & 2 & 6 & 8 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \\ &\rightsquigarrow \begin{pmatrix} 1 & 2 & 1 & 4 & 11 \\ 0 & 0 & 1 & 3 & 4 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1 & 2 & 1 & 4 & 0 \\ 0 & 0 & 1 & 3 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1 & 2 & 0 & 1 & 0 \\ 0 & 0 & 1 & 3 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \end{aligned}$$

Nach Satz IV.3.11(e) bilden die Vektoren

$$b_1 = \begin{pmatrix} -2 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \quad b_2 = \begin{pmatrix} -1 \\ 0 \\ -3 \\ 1 \\ 0 \end{pmatrix}$$

eine Basis von L und es gilt $\dim(L) = 2$. Insbesondere ist

$$\phi: \mathbb{R}^2 \xrightarrow{\cong} L, \quad \phi \begin{pmatrix} s \\ t \end{pmatrix} = s \begin{pmatrix} -2 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} + t \begin{pmatrix} -1 \\ 0 \\ -3 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} -2s-t \\ s \\ -3t \\ t \\ 0 \end{pmatrix},$$

ein Isomorphismus, d.h. eine Parameterdarstellung. Nach Satz IV.3.11(b) erhalten wir aus den nicht-trivialen Zeilen der reduzierten Zeilenstufenform folgendes

minimale Gleichungssystem für L :

$$\begin{array}{rcccc} x_1 & +2x_2 & & +x_4 & = & 0 \\ & & & & & \\ & & x_3 & +3x_4 & = & 0 \\ & & & & & \\ & & & & x_5 & = & 0 \end{array}$$

Nach Satz IV.3.11(c) bilden die Einheitsvektoren

$$e_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \quad e_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \quad e_5 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}$$

eine Basis eines zu L komplementären Teilraums L' , der auch durch das minimale Gleichungssystem

$$\begin{array}{r} x_2 = 0 \\ x_4 = 0 \end{array}$$

beschrieben werden kann. Um die letzte Frage zu beantworten bringen wir A auf Spaltenstufenform:

$$\begin{pmatrix} 1 & 2 & 1 & 4 & 11 \\ 2 & 4 & 4 & 14 & 30 \\ 3 & 6 & 3 & 12 & 34 \\ 2 & 4 & 6 & 20 & 41 \\ 1 & 2 & -1 & -2 & 7 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 2 & 0 & 2 & 6 & 8 \\ 3 & 0 & 0 & 0 & 1 \\ 2 & 0 & 4 & 12 & 19 \\ 1 & 0 & -2 & -6 & -4 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 2 & 0 & 2 & 6 & 8 \\ 3 & 0 & 0 & 0 & 1 \\ 2 & 0 & 4 & 12 & 19 \\ 1 & 0 & -2 & -6 & -4 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 2 & 2 & 0 & 0 & 0 \\ 3 & 0 & 1 & 0 & 0 \\ 2 & 4 & 3 & 0 & 0 \\ 1 & -2 & 4 & 0 & 0 \end{pmatrix}$$

Nach Satz IV.3.12(d) bildet die ersten drei Gleichungen von (IV.8),

$$\begin{array}{rcccccc} x_1 & +2x_2 & +x_3 & +4x_4 & +11x_5 & = & 0 \\ 2x_1 & +4x_2 & +4x_3 & +14x_4 & +30x_5 & = & 0 \\ 3x_1 & +6x_2 & +3x_3 & +12x_4 & +34x_5 & = & 0 \end{array}$$

ein minimales Gleichungssystem für L .

IV.3.20. BEISPIEL. Betrachte den Teilraum $W = \langle v_1, v_2, v_3, v_4, v_5 \rangle \subseteq \mathbb{R}^5$, der von den Vektoren

$$v_1 = \begin{pmatrix} 1 \\ 2 \\ 3 \\ 2 \\ 1 \end{pmatrix}, \quad v_2 = \begin{pmatrix} 2 \\ 4 \\ 6 \\ 4 \\ 2 \end{pmatrix}, \quad v_3 = \begin{pmatrix} 1 \\ 4 \\ 3 \\ 6 \\ -1 \end{pmatrix}, \quad v_4 = \begin{pmatrix} 4 \\ 14 \\ 12 \\ 20 \\ -2 \end{pmatrix}, \quad v_5 = \begin{pmatrix} 11 \\ 30 \\ 34 \\ 41 \\ 7 \end{pmatrix}$$

aufgespannt wird. Wir wollen eine Basis von W bestimmen, die aus gewissen der Vektoren v_i besteht. Mit Zeilenumformungen erhalten wir:

$$\begin{pmatrix} 1 & 2 & 1 & 4 & 11 \\ 2 & 4 & 4 & 14 & 30 \\ 3 & 6 & 3 & 12 & 34 \\ 2 & 4 & 6 & 20 & 41 \\ 1 & 2 & -1 & -2 & 7 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1 & 2 & 1 & 4 & 11 \\ 0 & 0 & 2 & 6 & 8 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 4 & 12 & 19 \\ 0 & 0 & -2 & -6 & -4 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1 & 2 & 1 & 4 & 11 \\ 0 & 0 & 2 & 6 & 8 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 3 & 3 \\ 0 & 0 & 0 & 0 & 4 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1 & 2 & 1 & 4 & 11 \\ 0 & 0 & 2 & 6 & 8 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Nach Satz IV.3.11(d) bilden die Vektoren v_1, v_3, v_5 eine Basis von W und es gilt $\dim(W) = 3$. Auch sehen wir nun, dass etwa v_1, v_2, v_5 keine Basis von W bildet.

IV.3.21. BEISPIEL. Betrachte den von den Vektoren

$$v_1 = \begin{pmatrix} 1 \\ 2 \\ 7 \\ 4 \\ 5 \\ 8 \end{pmatrix}, \quad v_2 = \begin{pmatrix} -1 \\ -2 \\ -6 \\ -1 \\ -1 \\ 1 \end{pmatrix}, \quad v_3 = \begin{pmatrix} 2 \\ 4 \\ 15 \\ 11 \\ 14 \\ 26 \end{pmatrix}, \quad v_4 = \begin{pmatrix} 3 \\ 6 \\ 20 \\ 9 \\ 11 \\ 17 \end{pmatrix},$$

aufgespannten Teilraum $W = \langle v_1, v_2, v_3, v_4 \rangle \subseteq \mathbb{R}^6$. Wir wollen einen zu W komplementären Teilraum W' bestimmen, genauer, eine Basis von W' angeben. Mittels Spaltenumformungen erhalten wir:

$$\begin{pmatrix} 1 & -1 & 2 & 3 \\ 2 & -2 & 4 & 6 \\ 7 & -6 & 15 & 20 \\ 4 & -1 & 11 & 9 \\ 5 & -1 & 14 & 11 \\ 8 & 1 & 26 & 17 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1 & 0 & 0 & 0 \\ 2 & 0 & 0 & 0 \\ 7 & 1 & 1 & -1 \\ 4 & 3 & 3 & -3 \\ 5 & 4 & 4 & -4 \\ 8 & 9 & 10 & -7 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1 & 0 & 0 & 0 \\ 2 & 0 & 0 & 0 \\ 7 & 1 & 0 & 0 \\ 4 & 3 & 0 & 0 \\ 5 & 4 & 0 & 0 \\ 8 & 9 & 1 & 2 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1 & 0 & 0 & 0 \\ 2 & 0 & 0 & 0 \\ 7 & 1 & 0 & 0 \\ 4 & 3 & 0 & 0 \\ 5 & 4 & 0 & 0 \\ 8 & 9 & 1 & 0 \end{pmatrix}$$

Nach Satz IV.3.12(c) spannen daher die Einheitsvektoren e_2, e_4, e_5 einen Teilraum $W' = \langle e_2, e_4, e_5 \rangle \subseteq \mathbb{R}^6$ auf, der zu W komplementär ist, $W \oplus W' = \mathbb{R}^6$. Beachte auch $\dim(W) = 3 = \dim(W')$.

IV.3.22. BEISPIEL. Wir wollen verifizieren, dass die Vektoren

$$v_1 = \begin{pmatrix} 2 \\ -2 \\ 4 \end{pmatrix}, \quad v_2 = \begin{pmatrix} 4 \\ -1 \\ 11 \end{pmatrix}, \quad v_3 = \begin{pmatrix} 5 \\ 3 \\ 18 \end{pmatrix}, \quad v_4 = \begin{pmatrix} 6 \\ 3 \\ 21 \end{pmatrix}, \quad v_5 = \begin{pmatrix} 7 \\ -6 \\ 19 \end{pmatrix}$$

ein Erzeugendensystem von \mathbb{R}^3 sind, und drei dieser Vektoren bestimmen, die eine Basis von \mathbb{R}^3 bilden. Wir fassen die Vektoren zu einer Matrix zusammen:

$$A = \begin{pmatrix} 2 & 4 & 5 & 6 & 7 \\ -2 & -1 & 3 & 3 & -6 \\ 4 & 11 & 18 & 21 & 19 \end{pmatrix}$$

Mittels Zeilenumformungen erhalten wir:

$$\begin{pmatrix} 2 & 4 & 5 & 6 & 7 \\ -2 & -1 & 3 & 3 & -6 \\ 4 & 11 & 18 & 21 & 19 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 2 & 4 & 5 & 6 & 7 \\ 0 & 3 & 8 & 9 & 1 \\ 0 & 3 & 8 & 9 & 5 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 2 & 4 & 5 & 6 & 7 \\ 0 & 3 & 8 & 9 & 1 \\ 0 & 0 & 0 & 0 & 4 \end{pmatrix}$$

Somit ist $\text{rank}(A) = 3$, also bilden die Vektoren v_1, \dots, v_5 ein Erzeugendensystem von \mathbb{R}^3 , siehe Korollar IV.3.4. Nach Satz IV.3.11(d) bilden die Vektoren v_1, v_2, v_5 eine Basis von \mathbb{R}^3 .

IV.3.23. BEISPIEL. Es bezeichne $W \subseteq \mathbb{R}^5$ den Teilraum aller y , für die das folgende Gleichungssystem lösbar ist:

$$\begin{array}{rcccccl} x_1 & & -x_3 & & +2x_5 & = & y_1 \\ x_1 & +x_2 & +x_3 & +3x_4 & +x_5 & = & y_2 \\ 7x_1 & +5x_2 & +3x_3 & +15x_4 & +9x_5 & = & y_3 \\ 5x_1 & +2x_2 & & +7x_4 & +12x_5 & = & y_4 \\ 17x_1 & +7x_2 & -x_3 & +23x_4 & +35x_5 & = & y_5 \end{array} \quad (\text{IV.9})$$

Wir wollen nun $\dim(W)$ und eine Basis von W bestimmen, und auch ein minimales Gleichungssystem für W angeben. Beachte, dass W genau der Spaltenraum der Koeffizientenmatrix

$$\begin{pmatrix} 1 & 0 & -1 & 0 & 2 \\ 1 & 1 & 1 & 3 & 1 \\ 7 & 5 & 3 & 15 & 9 \\ 5 & 2 & 0 & 7 & 12 \\ 17 & 7 & -1 & 23 & 35 \end{pmatrix}$$

ist. Durch Spaltenumformungen bringen wir diese Matrix auf reduzierte Spaltenstufenform:

$$\begin{pmatrix} 1 & 0 & -1 & 0 & 2 \\ 1 & 1 & 1 & 3 & 1 \\ 7 & 5 & 3 & 15 & 9 \\ 5 & 2 & 0 & 7 & 12 \\ 17 & 7 & -1 & 23 & 35 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 2 & 3 & -1 \\ 7 & 5 & 10 & 15 & -5 \\ 5 & 2 & 5 & 7 & 2 \\ 17 & 7 & 16 & 23 & 1 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 2 & 5 & 0 & 0 & 0 \\ 3 & 2 & 1 & 1 & 4 \\ 10 & 7 & 2 & 2 & 8 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 2 & 5 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 4 & 3 & 2 & 0 & 0 \end{pmatrix}$$

Daraus lesen wir $\dim(W) = 3$ ab und erhalten auch eine Basis

$$b_1 = \begin{pmatrix} 1 \\ 0 \\ 2 \\ 0 \\ 4 \end{pmatrix}, \quad b_2 = \begin{pmatrix} 0 \\ 1 \\ 5 \\ 0 \\ 3 \end{pmatrix}, \quad b_3 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 2 \end{pmatrix},$$

von W , vgl. Satz IV.3.12(b). Mit Satz IV.3.12(e) erhalten wir auch ein minimales Gleichungssystem für W :

$$\begin{aligned} -2y_1 & -5y_2 & +y_3 & & & = 0 \\ -4y_1 & -3y_2 & & -2y_4 & +y_5 & = 0 \end{aligned} \quad (\text{IV.10})$$

Das ursprüngliche Gleichungssystem (IV.9) ist also genau dann lösbar, wenn die rechte Seite den beiden Gleichungen (IV.10) genügt.

IV.4. Inhomogene Gleichungssysteme. Sei wieder $A \in M_{m \times n}(\mathbb{K})$ eine Matrix. Wir wollen nun, für fixes $y \in \mathbb{K}^m$, das inhomogene System $Ax = y$ lösen, d.h. alle $x \in \mathbb{K}^n$ bestimmen, für die $Ax = y$ gilt.

IV.4.1. SATZ. Sei $A \in M_{m \times n}(\mathbb{K})$ und $y \in \mathbb{K}^m$. Das inhomogene System

$$Ax = y \quad (\text{IV.11})$$

ist genau dann lösbar, wenn

$$\text{rank}(A) = \text{rank}(A|y)$$

gilt. Durch Zeilenumformungen lässt sich die erweiterte Matrix $(A|y)$ auf die Form $(\tilde{A}|\tilde{y})$ bringen, wobei \tilde{A} (reduzierte) Zeilenstufenform hat. Sind $1 \leq j_1 < j_2 < \dots < j_k \leq n$ wie in Definition IV.3.10, so ist (IV.11) also genau dann lösbar, wenn $\tilde{y}_{k+1} = \tilde{y}_{k+2} = \dots = \tilde{y}_m = 0$ gilt. In diesem Fall liefern die ersten k Zeilen von $(\tilde{A}|\tilde{y})$ ein minimales inhomogenes Gleichungssystem, das dieselbe Lösungsmenge wie (IV.11) besitzt. Ist darüber hinaus die Zeilenstufenform \tilde{A} reduziert, dann bildet

$$\xi = \tilde{y}_1 e_{j_1} + \dots + \tilde{y}_k e_{j_k}$$

eine spezielle Lösung von (IV.11), d.h. es gilt $A\xi = y$. Mit Satz IV.3.11(e) erhalten wir aus \tilde{A} eine Basis b_1, \dots, b_{n-k} für den Lösungsraum des homogenen Systems $Ax = 0$. Die allgemeine Lösung von (IV.11) ist dann von der Form

$$x = \xi + s_1 b_1 + \dots + s_{n-k} b_{n-k}, \quad s_1, \dots, s_{n-k} \in \mathbb{K}.$$

BEWEIS. Bezeichnen $A = (a_1 | \dots | a_n)$ die Spalten von A , dann gilt

$$\text{rank}(A) = \dim(\langle a_1, \dots, a_n \rangle) \quad \text{und} \quad \text{rank}(A|y) = \dim(\langle a_1, \dots, a_n, y \rangle).$$

Daraus lesen wir eine spezielle Lösung ξ sowie eine Basis b_1, b_2, b_3 für den Lösungsraum des homogenen System ab, vgl. Satz IV.4.1:

$$\xi = \begin{pmatrix} 3 \\ 2 \\ 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \quad b_1 = \begin{pmatrix} -5 \\ -1 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \quad b_2 = \begin{pmatrix} -6 \\ 2 \\ 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \quad b_3 = \begin{pmatrix} -1 \\ -2 \\ 0 \\ 0 \\ -3 \\ 1 \end{pmatrix}.$$

Die allgemeine Lösung des Gleichungssystems (IV.12) ist daher

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \end{pmatrix} = \begin{pmatrix} 3 \\ 2 \\ 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} + s_1 \begin{pmatrix} -5 \\ -1 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} + s_2 \begin{pmatrix} -6 \\ 2 \\ 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} + s_3 \begin{pmatrix} -1 \\ -2 \\ 0 \\ 0 \\ -3 \\ 1 \end{pmatrix}, \quad s_1, s_2, s_3 \in \mathbb{R}.$$

Auch ein minimales Gleichungssystem für den Lösungsraum lässt sich ablesen:

$$\begin{array}{cccccc} x_1 & & +5x_3 & +6x_4 & & +x_6 & = & 3 \\ & x_2 & & +x_3 & -2x_4 & & +2x_6 & = & 2 \\ & & & 3 & & x_5 & +3x_6 & = & 1 \end{array}$$

IV.4.3. BEISPIEL. Wir wollen alle Lösungen des linearen Gleichungssystems

$$\begin{array}{cccccc} x_1 & +2x_2 & +3x_3 & +4x_4 & = & 4 \\ -x_1 & +3x_2 & +3x_3 & +3x_4 & = & -1 \\ x_1 & +7x_2 & +17x_3 & +20x_4 & = & 9 \\ 2x_1 & -x_2 & -16x_3 & -17x_4 & = & 2 \end{array} \quad (\text{IV.13})$$

bestimmen. Durch Zeilenumformungen erhalten wir:

$$\begin{aligned} \left(\begin{array}{cccc|c} 1 & 2 & 3 & 4 & 4 \\ -1 & 3 & 3 & 3 & -1 \\ 1 & 7 & 17 & 20 & 9 \\ 2 & -1 & -16 & -17 & 2 \end{array} \right) & \rightsquigarrow \left(\begin{array}{cccc|c} 1 & 2 & 3 & 4 & 4 \\ 0 & 5 & 6 & 7 & 3 \\ 0 & 5 & 14 & 16 & 5 \\ 0 & -5 & -22 & -25 & -6 \end{array} \right) \\ & \rightsquigarrow \left(\begin{array}{cccc|c} 1 & 2 & 3 & 4 & 4 \\ 0 & 5 & 6 & 7 & 3 \\ 0 & 0 & 8 & 9 & 2 \\ 0 & 0 & -16 & -18 & -3 \end{array} \right) \rightsquigarrow \left(\begin{array}{cccc|c} 1 & 2 & 3 & 4 & 4 \\ 0 & 5 & 6 & 7 & 3 \\ 0 & 0 & 8 & 9 & 2 \\ 0 & 0 & 0 & 0 & 1 \end{array} \right) \end{aligned}$$

Das Gleichungssystem (IV.13) besitzt daher keine Lösung, vgl. Satz IV.4.1.

Unter einem *affinen Teilraum* eines Vektorraums V verstehen wir jede Teilmenge der Form

$$E = \xi + W = \{\xi + w \mid w \in W\},$$

wobei W einen Teilraum von V bezeichnet und $\xi \in V$. In diesem Fall ist

$$W \xrightarrow{\cong} E, \quad w \mapsto \xi + w,$$

eine Bijektion, aber i.A. keine lineare Abbildung. Unter der Dimension des affinen Teilraums E verstehen wir die Dimension des (dazu parallelen) Teilraums W . Ist

b_1, \dots, b_k eine Basis von W , dann ist

$$\phi: \mathbb{K}^k \xrightarrow{\cong} E, \quad \phi(s) = \xi + s_1 b_1 + \dots + s_k b_k,$$

eine Bijektion, d.h. eine Parametrisierung von E , aber i.A. nicht linear. Sei $m = \dim(V)$ und $\alpha_1, \dots, \alpha_{m-k}$ ein Gleichungssystem für W , d.h. $W = \bigcap_{i=1}^{m-k} \ker(\alpha_i)$. Dann stimmt E also mit der Lösungsmenge des folgenden inhomogenen Systems überein:

$$\begin{aligned} \alpha_1(v) &= \alpha_1(\xi) \\ &\vdots \\ \alpha_{m-k}(v) &= \alpha_{m-k}(\xi) \end{aligned}$$

Jeder affine Teilräume $E \subseteq V$ kann daher durch $\dim(V) - \dim(E)$ inhomogenen lineare Gleichungen beschrieben werden. Umgekehrt ist die Lösungsmenge eines inhomogenen Gleichungssystems stets ein affiner Teilraum, oder leer, siehe Satz IV.4.1. Soll ein Gleichungssystem für einen affinen Teilraum $E = \xi + W$ gefunden werden, so ist es zweckmäßig zunächst ein homogenes Gleichungssystem für den Teilraum W zu bestimmen, die Konstanten auf der rechten Seite lassen sich dann durch Einsetzen des Punktes ξ ermitteln.

IV.4.4. BEISPIEL. Wir wollen die Dimension sowie ein minimales Gleichungssystem des affinen Teilraums

$$E = \left(\begin{array}{c} -1 \\ 2 \\ 3 \\ 4 \end{array} \right) + \left\langle \left(\begin{array}{c} 1 \\ 0 \\ -1 \\ 1 \end{array} \right), \left(\begin{array}{c} 1 \\ 2 \\ 3 \\ 7 \end{array} \right), \left(\begin{array}{c} 3 \\ 2 \\ 1 \\ 9 \end{array} \right) \right\rangle \subseteq \mathbb{R}^4$$

bestimmen. Mittels Spaltenumformungen erhalten wir

$$\left(\begin{array}{ccc} 1 & 1 & 3 \\ 0 & 2 & 2 \\ -1 & 3 & 1 \\ 1 & 7 & 9 \end{array} \right) \rightsquigarrow \left(\begin{array}{ccc} 1 & 0 & 0 \\ 0 & 2 & 2 \\ -1 & 4 & 4 \\ 1 & 6 & 6 \end{array} \right) \rightsquigarrow \left(\begin{array}{ccc} 1 & 0 & 0 \\ 0 & 2 & 2 \\ -1 & 4 & 4 \\ 1 & 6 & 0 \end{array} \right) \rightsquigarrow \left(\begin{array}{ccc} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -1 & 2 & 0 \\ 1 & 3 & 0 \end{array} \right),$$

es gilt daher $\dim(E) = 2$. Daraus lesen wir zunächst ein minimales Gleichungssystem für den zu E parallelen Teilraum ab, siehe Satz IV.3.12(e):

$$\begin{array}{cccc} x_1 & -2x_2 & +x_3 & = 0 \\ -x_1 & -3x_2 & & +x_4 = 0 \end{array}$$

Einsetzen des Punktes liefert dann folgendes minimale Gleichungssystem für E :

$$\begin{array}{cccc} x_1 & -2x_2 & +x_3 & = -2 \\ -x_1 & -3x_2 & & +x_4 = -1 \end{array}$$

IV.5. Matrizeninversion. Wir wollen in diesem Abschnitt einen effizienten Algorithmus zur Berechnung der Inversen einer Matrix besprechen. Wir beginnen mit folgender Charakterisierungen der Invertierbarkeit einer Matrix, die wir sofort durch Kombination der Korollare IV.3.4 und IV.3.5 erhalten:

IV.5.1. KOROLLAR (Invertierbare Matrizen). *Sei \mathbb{K} ein Körper. Für eine quadratische Matrix $A \in M_{n \times n}(\mathbb{K})$ sind folgende Aussagen äquivalent:*

(a) *Die Matrix A ist invertierbar.*

- (b) Die lineare Abbildung $\psi: \mathbb{K}^n \rightarrow \mathbb{K}^n$, $\psi(x) = Ax$, ist ein Isomorphismus.
 (c) Das Gl.system $Ax = y$ hat für jedes $y \in \mathbb{K}^n$ genau eine Lösung $x \in \mathbb{K}^n$.
 (d) Die Spalten von A bilden eine Basis von \mathbb{K}^n .
 (e) Die Zeilen von A bilden eine Basis von $M_{1 \times n}(\mathbb{K}) \cong (\mathbb{K}^n)^* \cong \mathbb{K}^n$.
 (f) Es gilt $\text{rank}(A) = n$.

IV.5.2. BEISPIEL. Wir wollen verifizieren, dass die Vektoren

$$v_1 = \begin{pmatrix} 2 \\ -2 \\ 2 \\ -4 \end{pmatrix}, \quad v_2 = \begin{pmatrix} 1 \\ -2 \\ 2 \\ 4 \end{pmatrix}, \quad v_3 = \begin{pmatrix} 0 \\ 2 \\ 2 \\ 8 \end{pmatrix}, \quad v_4 = \begin{pmatrix} 6 \\ -5 \\ 7 \\ -3 \end{pmatrix}$$

eine Basis von \mathbb{R}^4 bilden. Nach Korollar IV.5.1 ist dies genau dann der Fall, wenn die Matrix

$$A = \begin{pmatrix} 2 & 1 & 0 & 6 \\ -2 & 2 & 2 & -5 \\ 2 & -2 & 2 & 7 \\ -4 & 4 & 8 & -3 \end{pmatrix}$$

Rang 4 hat. Mittels Zeilenumformungen erhalten wir

$$\begin{pmatrix} 2 & 1 & 0 & 6 \\ -2 & 2 & 2 & -5 \\ 2 & -2 & 2 & 7 \\ -4 & 4 & 8 & -3 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 2 & 1 & 0 & 6 \\ 0 & 3 & 2 & 1 \\ 0 & -3 & 2 & 1 \\ 0 & 6 & 8 & 9 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 2 & 1 & 0 & 6 \\ 0 & 3 & 2 & 1 \\ 0 & 0 & 4 & 2 \\ 0 & 0 & 4 & 7 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 2 & 1 & 0 & 6 \\ 0 & 3 & 2 & 1 \\ 0 & 0 & 4 & 2 \\ 0 & 0 & 0 & 5 \end{pmatrix}$$

also bilden die Vektoren v_1, v_2, v_3, v_4 tatsächlich eine Basis von \mathbb{R}^4 .

Wollen wir die Inverse einer Matrix $A \in M_{n \times n}(\mathbb{K})$ berechnen, dann müssen wir also jene Matrix $X \in M_{n \times n}(\mathbb{K})$ bestimmen, für die $AX = I_n$ gilt. Dies kann als lineares Gleichungssystem mit n^2 vielen Gleichungen in den n^2 vielen Einträgen von X verstanden werden. Allerdings zerfällt dieses Gleichungssystem in n unabhängige Systeme, eines für jede Spalte von X , mit je n Gleichungen in n Unbekannten,

$$A \begin{pmatrix} x_{11} \\ \vdots \\ x_{n1} \end{pmatrix} = e_1, \quad A \begin{pmatrix} x_{12} \\ \vdots \\ x_{n2} \end{pmatrix} = e_2, \quad \dots \quad A \begin{pmatrix} x_{1n} \\ \vdots \\ x_{nn} \end{pmatrix} = e_n.$$

Diese Systeme lassen sich bequem gleichzeitig mit folgendem Algorithmus lösen:

IV.5.3. SATZ (Algorithmus zur Bestimmung der Inversen). *Ist $A \in M_{n \times n}(\mathbb{K})$ eine invertierbare $(n \times n)$ -Matrix, dann kann die $n \times (2n)$ -Matrix $(A|I_n)$ durch Zeilenumformungen auf die Gestalt $(I_n|B)$ gebracht werden und es gilt $A^{-1} = B$.*

BEWEIS. Die Matrix A kann durch Zeilenumformungen auf reduzierte Zeilenstufenform gebracht werden, siehe Satz IV.3.11. Wegen der Invertierbarkeit von A ist $\text{rank}(A) = n$, also muss die reduzierte Zeilenstufenform mit der Einheitsmatrix I_n übereinstimmen. Nach Lemma IV.3.9 entsprechen diese Zeilenumformungen gerade einer Multiplikation von links, $A \rightsquigarrow TA$, mit einer invertierbaren Matrix $T \in \text{GL}_n(\mathbb{K})$. Es gilt daher $TA = I_n$, also $T = A^{-1}$. Wenden wir die selben Zeilenumformungen auf die erweiterte Matrix $(A|I_n)$ an, erhalten wir also $(A|I_n) \rightsquigarrow T(A|I_n) = (TA|TI_n) = (I_n|T) = (I_n|A^{-1})$. \square

IV.5.4. BEISPIEL. Wir wollen die Inverse der Matrix

$$A = \begin{pmatrix} 1 & -2 & 3 \\ -2 & 6 & -3 \\ -3 & 8 & -3 \end{pmatrix} \in M_{3 \times 3}(\mathbb{R})$$

bestimmen. Mittels Zeilenumformungen erhalten wir:

$$\begin{aligned} \left(\begin{array}{ccc|ccc} 1 & -2 & 3 & 1 & 0 & 0 \\ -2 & 6 & -3 & 0 & 1 & 0 \\ -3 & 8 & -3 & 0 & 0 & 1 \end{array} \right) &\rightsquigarrow \left(\begin{array}{ccc|ccc} 1 & -2 & 3 & 1 & 0 & 0 \\ 0 & 2 & 3 & 2 & 1 & 0 \\ 0 & 2 & 6 & 3 & 0 & 1 \end{array} \right) \\ &\rightsquigarrow \left(\begin{array}{ccc|ccc} 1 & -2 & 3 & 1 & 0 & 0 \\ 0 & 2 & 3 & 2 & 1 & 0 \\ 0 & 0 & 3 & 1 & -1 & 1 \end{array} \right) &\rightsquigarrow \left(\begin{array}{ccc|ccc} 1 & -2 & 0 & 0 & 1 & -1 \\ 0 & 2 & 0 & 1 & 2 & -1 \\ 0 & 0 & 3 & 1 & -1 & 1 \end{array} \right) \\ &\rightsquigarrow \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & 3 & -2 \\ 0 & 2 & 0 & 1 & 2 & -1 \\ 0 & 0 & 3 & 1 & -1 & 1 \end{array} \right) &\rightsquigarrow \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & 3 & -2 \\ 0 & 1 & 0 & 1/2 & 1 & -1/2 \\ 0 & 0 & 1 & 1/3 & -1/3 & 1/3 \end{array} \right) \end{aligned}$$

Die Inverse von A ist daher

$$A^{-1} = \begin{pmatrix} 1 & 3 & -2 \\ 1/2 & 1 & -1/2 \\ 1/3 & -1/3 & 1/3 \end{pmatrix} = \frac{1}{6} \begin{pmatrix} 6 & 18 & -12 \\ 3 & 6 & -3 \\ 2 & -2 & 2 \end{pmatrix}.$$

IV.5.5. BEISPIEL. Wir wollen die Inverse der komplexen Matrix

$$A = \begin{pmatrix} 1 & 1+i & 1-i \\ 2 & 2+i & 2+2i \\ 0 & 1-3i & -4+13i \end{pmatrix} \in M_{3 \times 3}(\mathbb{C})$$

bestimmen. Mittels Zeilenumformungen erhalten wir:

$$\begin{aligned} \left(\begin{array}{ccc|ccc} 1 & 1+i & 1-i & 1 & 0 & 0 \\ 2 & 2+i & 2+2i & 0 & 1 & 0 \\ 0 & 1-3i & -4+13i & 0 & 0 & 1 \end{array} \right) &\rightsquigarrow \left(\begin{array}{ccc|ccc} 1 & 1+i & 1-i & 1 & 0 & 0 \\ 0 & -i & 4i & -2 & 1 & 0 \\ 0 & 1-3i & -4+13i & 0 & 0 & 1 \end{array} \right) \\ &\rightsquigarrow \left(\begin{array}{ccc|ccc} 1 & 1+i & 1-i & 1 & 0 & 0 \\ 0 & 1 & -4 & -2i & i & 0 \\ 0 & 1-3i & -4+13i & 0 & 0 & 1 \end{array} \right) &\rightsquigarrow \left(\begin{array}{ccc|ccc} 1 & 0 & 5+3i & -1+2i & 1-i & 0 \\ 0 & 1 & -4 & -2i & i & 0 \\ 0 & 0 & i & 6+2i & -3-i & 1 \end{array} \right) \\ &\rightsquigarrow \left(\begin{array}{ccc|ccc} 1 & 0 & 5+3i & -1+2i & 1-i & 0 \\ 0 & 1 & -4 & -2i & i & 0 \\ 0 & 0 & 1 & 2-6i & -1+3i & -i \end{array} \right) \\ &\rightsquigarrow \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & -29+26i & 15-13i & -3+5i \\ 0 & 1 & 0 & 8-26i & -4+13i & -4i \\ 0 & 0 & 1 & 2-6i & -1+3i & -i \end{array} \right) \end{aligned}$$

Die Inverse von A ist daher:

$$A^{-1} = \begin{pmatrix} -29 + 26\mathbf{i} & 15 - 13\mathbf{i} & -3 + 5\mathbf{i} \\ 8 - 26\mathbf{i} & -4 + 13\mathbf{i} & -4\mathbf{i} \\ 2 - 6\mathbf{i} & -1 + 3\mathbf{i} & -\mathbf{i} \end{pmatrix}$$

IV.5.6. BEISPIEL. Wir wollen zeigen, dass das Gleichungssystem

$$\begin{aligned} 4x_2 - 4x_3 + 7x_4 &= y_1 \\ 2x_2 - 2x_3 + 2x_4 &= y_2 \\ x_1 + 5x_2 - 6x_3 + 7x_4 &= y_3 \\ x_1 + 3x_2 - 5x_3 + 4x_4 &= y_4 \end{aligned} \tag{IV.14}$$

für jede rechte Seite $y_1, y_2, y_3, y_4 \in \mathbb{R}$ eindeutig lösbar ist und diese Lösung bestimmen. Wir betrachten die Koeffizientenmatrix des Gleichungssystems,

$$A = \begin{pmatrix} 0 & 4 & -4 & 7 \\ 0 & 2 & -2 & 2 \\ 1 & 5 & -6 & 7 \\ 1 & 3 & -5 & 4 \end{pmatrix} \in M_{4 \times 4}(\mathbb{R}).$$

Mittels Zeilenumformungen erhalten wir:

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

Die Matrix A ist daher invertierbar mit Inverser

$$A^{-1} = \begin{pmatrix} -1 & -3/2 & 2 & -1 \\ -2/3 & 5/6 & 1 & -1 \\ -1/3 & -1/3 & 1 & -1 \\ 1/3 & -2/3 & 0 & 0 \end{pmatrix}.$$

Das Gleichungssysteme (IV.14) ist somit eindeutig lösbar und $x = A^{-1}y$ die gesuchte Lösung, d.h.

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} -1 & -3/2 & 2 & -1 \\ -2/3 & 5/6 & 1 & -1 \\ -1/3 & -1/3 & 1 & -1 \\ 1/3 & -2/3 & 0 & 0 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{pmatrix} = \begin{pmatrix} -y_1 - \frac{3}{2}y_2 + 2y_3 - y_4 \\ -\frac{2}{3}y_1 + \frac{5}{6}y_2 + y_3 - y_4 \\ -\frac{1}{3}y_1 - \frac{1}{3}y_2 + y_3 - y_4 \\ \frac{1}{3}y_1 - \frac{2}{3}y_2 \end{pmatrix}.$$

IV.5.7. BEISPIEL. Es soll die Inverse folgender Matrix über dem Körper $\mathbb{K} = \mathbb{Z}_2$ bestimmt werden:

$$A = \begin{pmatrix} 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 & 1 \end{pmatrix} \in M_{5 \times 5}(\mathbb{Z}_2)$$

Mittels Zeilenumformungen erhalten wir:

$$\begin{aligned} & \left(\begin{array}{ccccc|ccccc} 1 & 0 & 1 & 0 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 1 \end{array} \right) \rightsquigarrow \left(\begin{array}{ccccc|ccccc} 1 & 0 & 1 & 0 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 1 & 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 & 0 & 1 & 0 & 0 & 0 & 1 \end{array} \right) \rightsquigarrow \left(\begin{array}{ccccc|ccccc} 1 & 0 & 1 & 0 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 1 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 1 & 0 & 0 & 1 & 1 \end{array} \right) \\ & \rightsquigarrow \left(\begin{array}{ccccc|ccccc} 1 & 0 & 1 & 0 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 1 & 0 & 0 & 1 & 1 \end{array} \right) \rightsquigarrow \left(\begin{array}{ccccc|ccccc} 1 & 0 & 1 & 0 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 1 \end{array} \right) \\ & \rightsquigarrow \left(\begin{array}{ccccc|ccccc} 1 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 1 \end{array} \right) \rightsquigarrow \left(\begin{array}{ccccc|ccccc} 1 & 0 & 0 & 0 & 1 & 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 1 \end{array} \right) \end{aligned}$$

die Inverse von A ist daher:

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

IV.5.8. BEMERKUNG. Analog zur Formel für die Inverse einer (2×2) -Matrix, siehe Beispiel II.4.8, gibt es auch eine explizite Formel für die Inversion von

$(n \times n)$ -Matrizen. Für (3×3) -Matrizen sieht diese wie folgt aus:

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}^{-1} = \frac{1}{D} \begin{pmatrix} |a_{22} & a_{23}| & -|a_{21} & a_{23}| & |a_{21} & a_{22}| \\ -|a_{12} & a_{13}| & |a_{11} & a_{13}| & -|a_{11} & a_{12}| \\ |a_{32} & a_{33}| & -|a_{31} & a_{33}| & |a_{31} & a_{32}| \end{pmatrix}^t \quad (\text{IV.15})$$

wobei $\begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc$ und

$$D := a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{31}a_{22}a_{13} - a_{21}a_{12}a_{33} - a_{11}a_{32}a_{23}.$$

Dabei ist A genau dann invertierbar wenn $D \neq 0$, d.h. genau dann wenn die rechte Seite in Gleichung (IV.15) Sinn macht. Die analogen Formeln für größere n sind noch komplexer und erfordern mehr Rechenaufwand als der oben beschriebene Algorithmus mit Zeilenumformungen.

IV.6. Basisdarstellung. Wir haben in Abschnitt II.4 lineare Abbildungen $\mathbb{K}^n \rightarrow \mathbb{K}^m$ durch Matrizen beschrieben. Wollen wir lineare Abbildungen zwischen allgemeinen endlich-dimensionalen Vektorräumen, $\varphi: V \rightarrow W$, durch Matrizen darstellen, ist es notwendig geordnete Basen B und C der beiden Vektorräume V und W zu fixieren. Damit kann jeder lineare Abbildung φ eine Matrix $[\varphi]_{CB}$ zugeordnet werden, und wieder entspricht die Matrizenmultiplikation der Komposition von Abbildungen, siehe Satz IV.6.15 unten. Diese Matrix $[\varphi]_{CB}$ hängt von der Wahl der Basen B und C ab. Gehen wir zu anderen Basen, über transformiert sich die Matrixdarstellung mit sogenannten Basiswechsellmatrizen, siehe Korollar IV.6.21 unten. In diesem Zusammenhang ist es auch zweckmäßig die zu einer Basis von V duale Basis von V^* zu betrachten, und damit wollen wir diesen Abschnitt beginnen.

Sei V ein endlich-dimensionaler \mathbb{K} -Vektorraum, $\dim(V) = n$, und b_1, \dots, b_n eine geordnete Basis von V . Nach Proposition IV.1.2(d) existiert zu jedem $i \in \{1, \dots, n\}$ ein eindeutig bestimmtes lineares Funktional $b_i^*: V \rightarrow \mathbb{K}$, sodass

$$b_i^*(b_j) = \delta_{ij} = \begin{cases} 1 & \text{falls } i = j, \text{ und} \\ 0 & \text{falls } i \neq j. \end{cases} \quad (\text{IV.16})$$

Wir erhalten somit Elemente $b_1^*, \dots, b_n^* \in V^*$.

IV.6.1. LEMMA. *In dieser Situation ist b_1^*, \dots, b_n^* eine Basis von V^* .*

BEWEIS. Seien $\lambda_1, \dots, \lambda_n \in \mathbb{K}$, sodass

$$\lambda_1 b_1^* + \dots + \lambda_n b_n^* = 0.$$

Auswerten bei b_j liefert unter Verwendung von (IV.16)

$$0 = (\lambda_1 b_1^* + \dots + \lambda_n b_n^*)(b_j) = \lambda_1 b_1^*(b_j) + \dots + \lambda_n b_n^*(b_j) = \lambda_j,$$

für jedes $j = 1, \dots, n$. Dies zeigt, dass die Vektoren b_1^*, \dots, b_n^* linear unabhängig in V^* sind. Nach Korollar IV.1.17 müssen diese Vektoren eine Basis von V^* bilden, denn es gilt $\dim(V^*) = \dim(V) = n$, siehe Korollar IV.2.15. \square

IV.6.2. DEFINITION (Duale Basis). Ist V ein endlich-dimensionaler Vektorraum, $\dim(V) = n$, und b_1, \dots, b_n eine geordnete Basis von V , dann werden die durch (IV.16) eindeutig bestimmten Funktionale b_1^*, \dots, b_n^* als die zu b_1, \dots, b_n duale Basis von V^* bezeichnet, vgl. Lemma IV.6.1.

IV.6.3. BEMERKUNG. Sei b_1, \dots, b_n eine Basis von \mathbb{K}^n und $B := (b_1 | \dots | b_n)$ die Matrix mit Spaltenvektoren b_i . Weiters bezeichnen a_1, \dots, a_n die Zeilen der Inversen Matrix B^{-1} . Fassen wir $a_i \in M_{1 \times n}(\mathbb{K})$ als Element von $(\mathbb{K}^n)^*$ auf, $M_{1 \times n}(\mathbb{K}) \cong L(\mathbb{K}^n, \mathbb{K}) = (\mathbb{K}^n)^*$, dann ist a_1, \dots, a_n die zu b_1, \dots, b_n duale Basis, genauer $b_i^* = \psi_{a_i}$, d.h. $b_i^*: \mathbb{K}^n \rightarrow \mathbb{K}$, $b_i^*(x) = a_i x$, $x \in \mathbb{K}^n$, denn die Relation $B^{-1}B = I_n$ ist offensichtlich zu den Bedingungen (IV.16) äquivalent. Die duale Basis einer Basis von \mathbb{K}^n lässt sich daher durch Matrizeninversion berechnen.

IV.6.4. BEISPIEL. Es soll die zur Basis $b_1 = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$, $b_2 = \begin{pmatrix} 3 \\ 7 \end{pmatrix}$ von \mathbb{K}^2 duale Basis von $(\mathbb{K}^2)^*$ bestimmt werden. Für die Inverse der Matrix $B = \begin{pmatrix} 1 & 3 \\ 2 & 7 \end{pmatrix}$ gilt $B^{-1} = \begin{pmatrix} 7 & -3 \\ -2 & 1 \end{pmatrix}$. Für die duale Basis erhalten wir also $b_1^* = \psi_{(7, -3)}$ und $b_2^* = \psi_{(-2, 1)}$. Expliziter, $b_1^*: \mathbb{K}^2 \rightarrow \mathbb{K}$, $b_1^*\begin{pmatrix} x \\ y \end{pmatrix} = (7, -3)\begin{pmatrix} x \\ y \end{pmatrix} = 7x - 3y$, und $b_2^*: \mathbb{K}^2 \rightarrow \mathbb{K}$, $b_2^*\begin{pmatrix} x \\ y \end{pmatrix} = (-2, 1)\begin{pmatrix} x \\ y \end{pmatrix} = -2x + y$.

IV.6.5. BEISPIEL. Wir wollen die zur Basis $b_1 = \begin{pmatrix} 1 \\ -1 \\ 3 \end{pmatrix}$, $b_2 = \begin{pmatrix} -1 \\ -1 \\ -2 \end{pmatrix}$, $b_3 = \begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix}$ von \mathbb{K}^3 duale Basis von $(\mathbb{K}^3)^*$ bestimmen. Zeilenumformungen liefern:

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

Daraus schließen wir

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

für die duale Basis gilt daher $b_1^* = \psi_{(-2, 0, 1)}$, $b_2^* = \psi_{(-4, -1, 2)}$, $b_3^* = \psi_{(-1, -1, 1)}$, genauer $b_1^*\begin{pmatrix} x \\ y \\ z \end{pmatrix} = -2x + z$, $b_2^*\begin{pmatrix} x \\ y \\ z \end{pmatrix} = -4x - y + 2z$ und $b_3^*\begin{pmatrix} x \\ y \\ z \end{pmatrix} = -x - y + z$.

IV.6.6. BEISPIEL (Duale Basis der Standardbasis). Die zur Standardbasis e_1, \dots, e_n von \mathbb{K}^n duale Basis e_1^*, \dots, e_n^* von $(\mathbb{K}^n)^*$ wird durch die $(1 \times n)$ -Matrizen e_1^t, \dots, e_n^t repräsentiert, genauer $e_i^* = \psi_{e_i^t}$, d.h. $e_i^*: \mathbb{K}^n \rightarrow \mathbb{K}$ ist gerade die Projektion auf die i -te Koordinate, $e_i^*(x) = x_i$, $x \in \mathbb{K}^n$.

IV.6.7. LEMMA. Sei V ein n -dimensionaler Vektorraum, b_1, \dots, b_n eine Basis von V , b_1^*, \dots, b_n^* die dazu duale Basis von V^* und $b_1^{**}, \dots, b_n^{**}$ die zu letzterer duale Basis von V^{**} . Dann gilt

$$b_i^{**} = \iota(b_i), \quad 1 \leq i \leq n,$$

wobei $\iota: V \rightarrow V^{**}$ die natürliche Abbildung aus Proposition III.4.13 bezeichnet. In anderen Worten: $b_i^{**}(\alpha) = \alpha(b_i)$, für alle $\alpha \in V^*$.

BEWEIS. Für $1 \leq i, j \leq n$ gilt nach (IV.16)

$$\iota(b_i)(b_j^*) = b_j^*(b_i) = \delta_{ji} = \delta_{ij} = b_i^{**}(b_j^*).$$

Aus der Eindeutigkeitsaussage in Proposition IV.1.2(d) folgt daher $\iota(b_i) = b_i^{**}$, denn diese beiden Funktionale $V^* \rightarrow \mathbb{K}$ stimmen auf der Basis b_1^*, \dots, b_n^* von V^* überein. \square

IV.6.8. BEMERKUNG. Sei V ein endlich-dimensionaler Vektorraum mit Basis b_1, \dots, b_n . Bezeichnet b_1^*, \dots, b_n^* die dazu duale Basis von V^* , dann existiert genau eine lineare Abbildung $\varphi: V \rightarrow V^*$ mit $\varphi(b_i) = b_i^*$, $i = 1, \dots, n$, und diese Abbildung ist ein Isomorphismus. Bezeichnet $b_1^{**}, \dots, b_n^{**}$ die zu b_1^*, \dots, b_n^* duale Basis von V^{**} , dann existiert analog ein Isomorphismus $\psi: V^* \rightarrow V^{**}$ mit $\psi(b_i^*) = b_i^{**}$, $i = 1, \dots, n$. Obwohl φ und ψ beide von der ursprünglichen Basis abhängen, ist deren Komposition unabhängig von dieser Basis, denn nach Lemma IV.6.7 gilt $\psi \circ \varphi = \iota: V \rightarrow V^{**}$.

IV.6.9. PROPOSITION. *Sei V ein endlich-dimensionaler Vektorraum. Jede Basis von V^* ist die duale Basis einer Basis von V .*

BEWEIS. Sei also β_1, \dots, β_n eine Basis von V^* , wobei $n = \dim(V^*)$. Weiters bezeichne $\beta_1^*, \dots, \beta_n^*$ die dazu duale Basis von V^{**} . Nach Korollar IV.2.15 ist $\iota: V \rightarrow V^{**}$ ein Isomorphismus. Setzen wir $b_j := \iota^{-1}(\beta_j^*)$, $1 \leq j \leq n$, dann bildet also b_1, \dots, b_n eine Basis von V , siehe Proposition IV.1.3. Bezeichnet b_1^*, \dots, b_n^* die dazu dual Basis von V^* , dann gilt für alle $1 \leq i, j \leq n$,

$$\beta_i(b_j) = \iota(b_j)(\beta_i) = \beta_j^*(\beta_i) = \delta_{ji} = \delta_{ij} = b_i^*(b_j),$$

also stimmen die Funktionale β_i und b_i^* auf einer Basis von V überein. Aus der Eindeutigkeitsaussage in Proposition IV.1.2(d) folgt daher $\beta_i = b_i^*$. Dies zeigt, dass β_1, \dots, β_n die zu b_1, \dots, b_n duale Basis ist. \square

Sei V ein n -dimensionaler Vektorraum über \mathbb{K} und $B = (b_1, \dots, b_n)$ eine geordnete Basis von V . Nach Proposition IV.1.2 ist

$$\phi_B: \mathbb{K}^n \xrightarrow{\cong} V, \quad \phi_B(x) := x_1 b_1 + \dots + x_n b_n,$$

ein linearer Isomorphismus. Zu jedem $v \in V$ existieren daher eindeutig bestimmte Skalare $x_1, \dots, x_n \in \mathbb{K}$, sodass $v = x_1 b_1 + \dots + x_n b_n$. Der Vektor

$$[v]_B := \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{K}^n$$

wird als *Koordinatenvektor von v bezüglich der Basis B* bezeichnet. Nach Konstruktion ist $V \rightarrow \mathbb{K}^n$, $v \mapsto [v]_B$, gerade die Umkehrabbildung von ϕ_B , für jedes $v \in V$ gilt daher

$$v = \phi_B([v]_B) = ([v]_B)_1 b_1 + \dots + ([v]_B)_n b_n, \quad (\text{IV.17})$$

und auch $[\phi_B(x)]_B = x$ für alle $x \in \mathbb{K}^n$. Für die Basisvektoren erhalten wir

$$[b_i]_B = e_i, \quad 1 \leq i \leq n.$$

Aus der Linearität der Umkehrabbildung folgt weiters

$$[v_1 + v_2]_B = [v_1]_B + [v_2]_B \quad \text{und} \quad [\lambda v]_B = \lambda[v]_B \quad (\text{IV.18})$$

für alle $v, v_1, v_2 \in V$ und $\lambda \in \mathbb{K}$. Die Komponenten von $[v]_B$ können auch mit Hilfe der zu B dualen Basis $B^* = (b_1^*, \dots, b_n^*)$ beschrieben werden, für die i -te Komponente gilt

$$([v]_B)_i = b_i^*(v), \quad 1 \leq i \leq n, \quad (\text{IV.19})$$

denn $b_i^*(v) = b_i^*(x_1 b_1 + \dots + x_n b_n) = x_1 b_i^*(b_1) + \dots + x_n b_i^*(b_n) = x_i = ([v]_B)_i$.

IV.6.10. BEMERKUNG. Ist etwa $B = (b_1, \dots, b_n)$ eine Basis von \mathbb{K}^n und sollen die Koordinaten eines Vektors $v \in \mathbb{K}^n$ bezüglich B bestimmt werden, so muss das Gleichungssystem

$$x_1 b_1 + \dots + x_n b_n = v$$

gelöst werden, die Koordinaten von v bezüglich B sind dann

$$[v]_B = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}.$$

Werden die Koordinaten von vielen verschiedenen Vektoren benötigt, dann ist es zweckmäßig zunächst die duale Basis B^* zu bestimmen, wir erhalten die Koordinaten eines beliebigen Vektors v dann sofort aus (IV.19),

$$[v]_B = \begin{pmatrix} b_1^*(v) \\ \vdots \\ b_n^*(v) \end{pmatrix} = (b_1 | \dots | b_n)^{-1} v,$$

vgl. Bemerkung IV.6.3.

IV.6.11. BEISPIEL. Betrachte die Basis $B = (b_1, b_2)$ von \mathbb{K}^2 , wobei $b_1 = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$ und $b_2 = \begin{pmatrix} 3 \\ 4 \end{pmatrix}$. Es sollen die Koordinaten des Vektors $v = \begin{pmatrix} 3 \\ 4 \end{pmatrix}$ bezüglich B bestimmt werden. In Beispiel IV.6.4 haben wir die zu B duale Basis berechnet, $b_1^* = \psi_{(7,-3)}$ und $b_2^* = \psi_{(-2,1)}$. Mit (IV.19) folgt $[v]_B = \begin{pmatrix} b_1^*(v) \\ b_2^*(v) \end{pmatrix} = \begin{pmatrix} 9 \\ -2 \end{pmatrix}$. Die Entwicklung des Vektors v in der Basis B ist daher $v = \begin{pmatrix} 3 \\ 4 \end{pmatrix} = 9b_1 - 2b_2 = 9 \begin{pmatrix} 1 \\ 2 \end{pmatrix} - 2 \begin{pmatrix} 3 \\ 4 \end{pmatrix}$.

IV.6.12. BEISPIEL. Betrachte die Basis $B = (b_1, b_2, b_3)$ von \mathbb{K}^3 , wobei

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

Es sollen die Koordinaten des Vektors $v = \begin{pmatrix} 3 \\ 2 \\ 1 \end{pmatrix} \in \mathbb{K}^3$ bezüglich der Basis B bestimmt werden. In Beispiel IV.6.5 haben wir die zu B duale Basis bestimmt,

$b_1^* = \psi_{(-2,0,1)}$, $b_2^* = \psi_{(-4,-1,2)}$ und $b_3^* = \psi_{(-1,-1,1)}$. Mit (IV.19) folgt

$$[v]_B = \begin{pmatrix} b_1^*(v) \\ b_2^*(v) \\ b_3^*(v) \end{pmatrix} = \begin{pmatrix} -5 \\ -12 \\ -4 \end{pmatrix}$$

die Entwicklung des Vektors v in der Basis B ist daher

$$v = \begin{pmatrix} 3 \\ 2 \\ 1 \end{pmatrix} = -5b_1 - 12b_2 - 4b_3 = -5 \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} - 12 \begin{pmatrix} -1 \\ -1 \\ -2 \end{pmatrix} - 4 \begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix}.$$

IV.6.13. BEISPIEL. Für die Koordinaten eines Vektors $x \in \mathbb{K}^n$ bezüglich der Standardbasis $E = (e_1, \dots, e_n)$ von \mathbb{K}^n gilt offensichtlich

$$[x]_E = x.$$

Nach Satz II.4.4 kann jedes lineare Funktional $\alpha \in (\mathbb{K}^n)^*$ durch einen Zeilenvektor $a = (a_1, \dots, a_n) \in M_{1 \times n}(\mathbb{K})$ beschrieben werden, $\alpha = \psi_a$, d.h. $\alpha(x) = ax$, $x \in \mathbb{K}^n$. Bezeichnet E^* die zur Standardbasis duale Basis von $(\mathbb{K}^n)^*$, dann gilt

$$[\alpha]_{E^*} = a^t,$$

denn $\alpha = a_1e_1^* + \dots + a_n e_n^*$, da ja beide Seiten auf den Basisvektoren e_i übereinstimmen, $\alpha(e_i) = \psi_a(e_i) = ae_i = a_i = (a_1e_1^* + \dots + a_n e_n^*)(e_i)$.

IV.6.14. BEISPIEL. Betrachte die Basis $B = (1, z, z^2)$ des Vektorraums $\mathbb{K}[z]_{\leq 2}$. Der Koordinatenvektor eines Polynoms $p = p_0 + p_1z + p_2z^2 \in \mathbb{K}[z]_{\leq 2}$ bezüglich B ist dann

$$[p]_B = [p_0 + p_1z + p_2z^2]_B = \begin{pmatrix} p_0 \\ p_1 \\ p_2 \end{pmatrix} \in \mathbb{K}^3.$$

Der i -te Vektor der dualen Basis ist jenes lineare Funktional $\mathbb{K}[z]_{\leq 2} \rightarrow \mathbb{K}$, das einem Polynom seinen i -ten Koeffizienten zuordnet.

Sei nun $\varphi: V \rightarrow W$ eine lineare Abbildung zwischen endlich-dimensionalen \mathbb{K} -Vektorräumen, $\dim(V) = n$ und $\dim(W) = m$. Weiters seien $B = (b_1, \dots, b_n)$ und $C = (c_1, \dots, c_m)$ geordnete Basen von V und W . Nach Satz II.4.4 ist die Komposition $\phi_C^{-1} \circ \varphi \circ \phi_B: \mathbb{K}^n \rightarrow \mathbb{K}^m$ durch Multiplikation mit einer eindeutig bestimmten $(m \times n)$ -Matrix gegeben. Diese Matrix wird mit $[\varphi]_{CB} \in M_{m \times n}(\mathbb{K})$ bezeichnet und die *Matrix der linearen Abbildung φ bezüglich der Basen B und C* genannt. Die Definition lässt sich übersichtlich in folgendem kommutativen Diagramm veranschaulichen:

$$\begin{array}{ccc} \mathbb{K}^n & \xrightarrow{\phi_B} & V \\ \downarrow [\varphi]_{CB} & \cong & \downarrow \varphi \\ \mathbb{K}^m & \xleftarrow{\phi_C^{-1}} & W \end{array} \quad \text{d.h.} \quad (\phi_C^{-1} \circ \varphi \circ \phi_B)(x) = [\varphi]_{CB} x, \quad x \in \mathbb{K}^n. \quad (\text{IV.20})$$

Da $[\varphi]_{CB} e_j$ mit der j -ten Spalte von $[\varphi]_{CB}$ übereinstimmt, ist letztere also durch $(\phi_C^{-1} \circ \varphi \circ \phi_B)(e_j) = \phi_C^{-1}(\varphi(\phi_B(e_j))) = \phi_C^{-1}(\varphi(b_j)) = [\varphi(b_j)]_C$ gegeben. In anderen Worten, wir erhalten den j -ten Spaltenvektor von $[\varphi]_{CB}$ aus den Koordinaten

bezüglich der Basis C des Bildes des j -ten Basisvektors von B , d.h. die Eintragungen in der j -ten Spalten sind gerade die Koordinaten von $\varphi(b_j)$ bezüglich C . Für die Eintragung in der i -ten Zeile der j -ten Spalte erhalten wir aus (IV.19)

$$([\varphi]_{CB})_{ij} = ([\varphi(b_j)]_C)_i = c_i^*(\varphi(b_j)), \quad 1 \leq i \leq m, \quad 1 \leq j \leq n, \quad (\text{IV.21})$$

wobei $C^* = (c_1^*, \dots, c_m^*)$ die zu C duale Basis von W^* bezeichnet.

IV.6.15. SATZ (Matrixdarstellung linearer Abbildungen). *Seien V, W und U drei endlich-dimensionale Vektorräume über \mathbb{K} . Weiters seien B, C und D geordnete Basen von V, W und U . Für jede lineare Abbildung $\varphi: V \rightarrow W$ und jedes $v \in V$ gilt dann*

$$[\varphi(v)]_C = [\varphi]_{CB}[v]_B, \quad (\text{IV.22})$$

die Koordinaten des Bildes $\varphi(v)$ können daher durch Multiplikation mit der Matrix $[\varphi]_{CB}$ aus den Koordinaten von v berechnet werden. Die Zuordnung

$$L(V, W) \cong M_{m \times n}(\mathbb{K}), \quad \varphi \leftrightarrow [\varphi]_{CB}, \quad (\text{IV.23})$$

ist ein linearer Isomorphismus, wobei $n = \dim(V)$ und $m = \dim(W)$. Für beliebige $\varphi, \varphi_1, \varphi_2 \in L(V, W)$ und $\lambda \in \mathbb{K}$ gilt daher

$$[\varphi_1 + \varphi_2]_{CB} = [\varphi_1]_{CB} + [\varphi_2]_{CB} \quad \text{und} \quad [\lambda\varphi]_{CB} = \lambda[\varphi]_{CB}.$$

Für jede weitere lineare Abbildung $\psi: W \rightarrow U$ haben wir

$$[\psi \circ \varphi]_{DB} = [\psi]_{DC}[\varphi]_{CB} \quad \text{sowie} \quad [\text{id}_V]_{BB} = I_n. \quad (\text{IV.24})$$

Die lineare Abbildung φ ist genau dann invertierbar, wenn die Matrix $[\varphi]_{BC}$ invertierbar ist, in diesem Fall gilt

$$[\varphi^{-1}]_{BC} = [\varphi]_{CB}^{-1}. \quad (\text{IV.25})$$

Für jede lineare Abbildung $\varphi: V \rightarrow W$ gilt weiters

$$\text{rank}(\varphi) = \text{rank}([\varphi]_{CB}). \quad (\text{IV.26})$$

Bezeichnen B^* und C^* die zu B bzw. C dualen Basen von V^* und W^* , dann gilt für die Matrix der dualen Abbildung $\varphi^t: W^* \rightarrow V^*$

$$[\varphi^t]_{B^*C^*} = [\varphi]_{CB}^t. \quad (\text{IV.27})$$

BEWEIS. Aus (IV.20) erhalten wir mit $x = [v]_B$ sofort

$$[\varphi]_{CB}[v]_B = \phi_C^{-1}(\varphi(\phi_B([v]_B))) = \phi_C^{-1}(\varphi(v)) = [\varphi(v)]_C,$$

also (IV.22). Beachte, dass

$$L(V, W) \cong L(\mathbb{K}^n, \mathbb{K}^m), \quad \varphi \leftrightarrow \phi_C^{-1} \circ \varphi \circ \phi_B,$$

ein linearer Isomorphismus mit Umkehrabbildung $\phi_C \circ \sigma \circ \phi_B^{-1} \leftrightarrow \sigma$ ist, denn

$$\phi_C^{-1} \circ (\varphi_1 + \varphi_2) \circ \phi_B = \phi_C^{-1} \circ \varphi_1 \circ \phi_B + \phi_C^{-1} \circ \varphi_2 \circ \phi_B$$

und $\phi_C^{-1} \circ (\lambda\varphi) \circ \phi_B = \lambda(\phi_C^{-1} \circ \varphi \circ \phi_B)$ für alle $\varphi, \varphi_1, \varphi_2 \in L(V, W)$ und $\lambda \in \mathbb{K}$, siehe Proposition II.3.18. Zusammen mit Satz II.4.4 folgt, dass (IV.23) einen linearen Isomorphismus bildet. Aus

$$\phi_D^{-1} \circ (\psi \circ \varphi) \circ \phi_B = (\phi_D^{-1} \circ \psi \circ \phi_C) \circ (\phi_C^{-1} \circ \varphi \circ \phi), \quad \phi_B^{-1} \circ \text{id}_V \circ \phi_B = \text{id}_{\mathbb{K}^n},$$

und Satz II.4.4 erhalten wir nun auch (IV.24). Beachte, dass φ genau dann invertierbar ist, wenn $\phi_C^{-1} \circ \varphi \circ \phi_B$ invertierbar ist, und dass in diesem Fall

$$\phi_B^{-1} \circ \varphi^{-1} \circ \phi_C = (\phi_C^{-1} \circ \varphi \circ \phi_B)^{-1}$$

gilt. Zusammen mit Satz II.4.4 sehen wir daher, dass φ genau dann invertierbar ist, wenn $[\varphi]_{CB}$ invertierbar ist, und dass in diesem Fall (IV.25) gilt.

Nach Definition des Rangs einer Matrix gilt $\text{rank}([\varphi]_{CB}) = \text{rank}(\phi_C^{-1} \circ \varphi \circ \phi_B)$, die Gleichung (IV.26) folgt daher aus Proposition IV.2.23.

Seien nun $B^* = (b_1^*, \dots, b_n^*)$ und $C^* = (c_1^*, \dots, c_m^*)$ die zu $B = (b_1, \dots, b_n)$ bzw. $C = (c_1, \dots, c_m)$ dualen Basen von V^* und W^* . Mit (IV.21) erhalten wir

$$\begin{aligned} ([\varphi^t]_{B^*C^*})_{ij} &= b_i^{**}(\varphi^t(c_j^*)) = b_i^{**}(c_j^* \circ \varphi) = \iota_V(b_i)(c_j^* \circ \varphi) \\ &= (c_j^* \circ \varphi)(b_i) = c_j^*(\varphi(b_i)) = ([\varphi]_{CB})_{ji} = ([\varphi]_{CB}^t)_{ij}, \end{aligned}$$

wobei $B^{**} = (b_1^{**}, \dots, b_n^{**})$ die zu B^* duale Basis von V^{**} bezeichnet, für die ja $\iota_V(b_i) = b_i^{**}$ gilt, siehe Lemma IV.6.7. Dies zeigt (IV.27). \square

IV.6.16. BEMERKUNG. Bezeichnen $B = (e_1, \dots, e_n)$ und $C = (e_1, \dots, e_m)$ die Standardbasen von \mathbb{K}^n und \mathbb{K}^m , dann stimmt der Isomorphismus aus Satz IV.6.15, $L(\mathbb{K}^n, \mathbb{K}^m) \cong M_{m \times n}(\mathbb{K})$, $\varphi \leftrightarrow [\varphi]_{CB}$, mit dem in Satz II.4.4 besprochenen Isomorphismus überein. In diesem Fall gilt nämlich $\phi_B = \text{id}_{\mathbb{K}^n}$ und $\phi_C = \text{id}_{\mathbb{K}^m}$. Die Matrix einer linearen Abbildung $\varphi: \mathbb{K}^n \rightarrow \mathbb{K}^m$ bezüglich der Standardbasen von \mathbb{K}^n und \mathbb{K}^m stimmt also mit der in Satz II.4.4 besprochenen Matrix überein.

IV.6.17. BEMERKUNG. Sei V ein endlich-dimensionaler Vektorraum über \mathbb{K} , $\dim(V) = n$, und B eine geordnete Basis von V . Dann ist die Zuordnung

$$\text{end}(V) \cong M_{n \times n}(\mathbb{K}), \quad \varphi \leftrightarrow [\varphi]_{BB}$$

ein Isomorphismus von \mathbb{K} -Algebren, der sich zu einem Gruppenisomorphismus

$$\text{GL}(V) \cong \text{GL}_n(\mathbb{K}), \quad \varphi \leftrightarrow [\varphi]_{BB}$$

einschränkt. Dies folgt sofort aus Satz IV.6.15.

IV.6.18. BEISPIEL. Sei V ein n -dimensionaler \mathbb{K} -Vektorraum und $\pi: V \rightarrow V$ ein Projektor. Sei b_1, \dots, b_k eine Basis von $\text{img}(\pi)$ und b_{k+1}, \dots, b_n eine Basis von

$\ker(\pi)$. Da $V = \text{img}(\pi) \oplus \ker(\pi)$ ist $B := (b_1, \dots, b_n)$ eine Basis von V und

$$[\pi]_{BB} = \begin{pmatrix} 1 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & & \vdots \\ 0 & \cdots & 1 & 0 & \cdots & 0 \\ 0 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & & \vdots & \vdots & & \vdots \\ 0 & \cdots & 0 & 0 & \cdots & 0 \end{pmatrix}, \quad [\pi']_{BB} = \begin{pmatrix} 0 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & & \vdots \\ 0 & \cdots & 0 & 0 & \cdots & 0 \\ 0 & \cdots & 0 & 1 & \cdots & 0 \\ \vdots & & \vdots & \vdots & & \vdots \\ 0 & \cdots & 0 & 0 & \cdots & 1 \end{pmatrix},$$

wobei $\pi' = \text{id}_V - \pi$ den komplementären Projektor auf $\ker(\pi)$ längs $\text{img}(\pi)$ bezeichnet. Für die Spiegelung $\sigma = \pi - \pi'$ an $\text{img}(\pi)$ längs $\ker(\pi)$ und die dazu komplementäre Spiegelung $\sigma' = -\sigma = \pi' - \pi$ an $\ker(\pi)$ längs $\text{img}(\pi)$ haben wir

$$[\sigma]_{BB} = \begin{pmatrix} 1 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & & \vdots \\ 0 & \cdots & 1 & 0 & \cdots & 0 \\ 0 & \cdots & 0 & -1 & \cdots & 0 \\ \vdots & & \vdots & \vdots & & \vdots \\ 0 & \cdots & 0 & 0 & \cdots & -1 \end{pmatrix}, \quad [\sigma']_{BB} = \begin{pmatrix} -1 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & & \vdots \\ 0 & \cdots & -1 & 0 & \cdots & 0 \\ 0 & \cdots & 0 & 1 & \cdots & 0 \\ \vdots & & \vdots & \vdots & & \vdots \\ 0 & \cdots & 0 & 0 & \cdots & 1 \end{pmatrix}.$$

IV.6.19. BEISPIEL. Wir betrachten den reellen Vektorraum der glatten Funktionen, $C^\infty(\mathbb{R}, \mathbb{R})$. Die Ableitung definiert eine lineare Abbildung

$$C^\infty(\mathbb{R}, \mathbb{R}) \rightarrow (\mathbb{R}, \mathbb{R}), \quad f \mapsto f',$$

denn $(f + g)' = f' + g'$ und $(\lambda f)' = \lambda f'$, für alle $f, g \in C^\infty(\mathbb{R}, \mathbb{R})$ und $\lambda \in \mathbb{R}$. Nach Beispiel III.2.22 sind die beiden Funktionen $\sin, \cos \in C^\infty(\mathbb{R}, \mathbb{R})$ linear unabhängig, spannen also einen 2-dimensionalen Teilraum $W := \langle \sin, \cos \rangle \subseteq C^\infty(\mathbb{R}, \mathbb{R})$ auf. Da $\sin' = \cos$ und $\cos' = -\sin$, schränkt sich die Ableitung zu einer linearen Abbildung $D: W \rightarrow W$, $D(f) := f'$, ein. Für die Matrix von D bezüglich der geordneten Basis $B = (\sin, \cos)$ von W erhalten wir

$$[D]_{BB} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}.$$

Mit (IV.24) folgt daraus etwa

$$[D \circ D]_{BB} = [D]_{BB}[D]_{BB} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix},$$

was genau $(D \circ D)(\sin) = \sin'' = -\sin$ und $(D \circ D)(\cos) = \cos'' = -\cos$ entspricht.

IV.6.20. BEISPIEL. Betrachte den $(n + 1)$ -dimensionalen Vektorraum $V = \mathbb{R}[z]_{\leq n}$ mit geordneter Basis $B = (1, z, z^2, z^3, \dots, z^n)$, siehe Beispiel III.3.10. Die Ableitung definiert eine lineare Abbildung $D: W \rightarrow W$, $D(p) := p'$, denn es gilt $(p + q)' = p' + q'$ und $(\lambda p)' = \lambda p'$, für alle $p, q \in \mathbb{R}[z]$ und $\lambda \in \mathbb{R}$. Da $D(z^k) = kz^{k-1}$,

ist die Matrix von D bezüglich der Basis B durch

$$[D]_{BB} = \begin{pmatrix} 0 & 1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 2 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 0 & 3 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & n-1 & 0 \\ 0 & 0 & 0 & 0 & \cdots & 0 & n \\ 0 & 0 & 0 & 0 & \cdots & 0 & 0 \end{pmatrix} \in M_{(n+1) \times (n+1)}(\mathbb{R}).$$

gegeben. Setzen wir $n = 3$ und betrachten die lineare Abbildung

$$P: \mathbb{R}[z]_{\leq 3} \rightarrow \mathbb{R}[z]_{\leq 3}, \quad P(p) := 4p + 3p' + 2p'' = (4\text{id} + 3D + 2D^2)(p),$$

dann erhalten wir mit den Rechenregeln aus Satz IV.6.15 sofort

$$\begin{aligned} [P]_{BB} &= 4I_n + 3[D]_{BB} + 2([D]_{BB})^2 \\ &= 4 \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} + 3 \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 \end{pmatrix} + 2 \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 \end{pmatrix}^2 = \begin{pmatrix} 4 & 3 & 4 & 0 \\ 0 & 4 & 6 & 12 \\ 0 & 0 & 4 & 9 \\ 0 & 0 & 0 & 4 \end{pmatrix} \end{aligned}$$

Da diese Matrix invertierbar ist, gibt es also zu jedem Polynom $q \in \mathbb{R}[z]_{\leq 3}$ genau ein Polynom $p \in \mathbb{R}[z]_{\leq 3}$ für das $4p + 3p' + 2p'' = q$ gilt. Durch Inversion der Matrix $[P]_{BB}$ ließe sich dieses Polynom p auch sofort berechnen.

Seien nun $B = (b_1, \dots, b_n)$ und $\tilde{B} = (\tilde{b}_1, \dots, \tilde{b}_n)$ zwei geordnete Basen von V . Unter der *Matrix zum Basiswechsel von B nach \tilde{B}* verstehen wir die Matrix

$$T_{\tilde{B}B} := [\text{id}_V]_{\tilde{B}B} \in M_{n \times n}(\mathbb{K}).$$

Nach Definition stimmt der j -te Spaltenvektor von $T_{\tilde{B}B}$ mit $[b_j]_{\tilde{B}}$ überein. Für den Eintrag in der i -ten Zeile der j -ten Spalte gilt daher

$$(T_{\tilde{B}B})_{ij} = ([b_j]_{\tilde{B}})_i = \tilde{b}_i^*(b_j), \quad 1 \leq i, j \leq n,$$

wobei $\tilde{B}^* = (\tilde{b}_1^*, \dots, \tilde{b}_n^*)$ die zu \tilde{B} duale Basis bezeichnet.

IV.6.21. KOROLLAR (Basiswechsel). *Seien V und W zwei endlich-dimensionale Vektorräume über \mathbb{K} . Weiters seien B und \tilde{B} zwei geordnete Basen von V und C und \tilde{C} zwei geordnete Basen von W . Die Basiswechselmatrizen sind stets invertierbar mit Inverser*

$$T_{\tilde{B}B}^{-1} = T_{B\tilde{B}} \tag{IV.28}$$

Für jedes $v \in V$ gilt

$$[v]_{\tilde{B}} = T_{\tilde{B}B}[v]_B, \tag{IV.29}$$

wir erhalten also die Koordinaten von v bezüglich der Basis \tilde{B} durch Multiplikation mit der Matrix $T_{\tilde{B}B}$ aus den Koordinaten bezüglich B . Für jede Abbildung $\varphi: V \rightarrow W$ gilt

$$[\varphi]_{\tilde{C}\tilde{B}} = T_{\tilde{C}C}[\varphi]_{CB}T_{B\tilde{B}}. \tag{IV.30}$$

Für die Basiswechselmatrix der dualen Basen gilt

$$T_{B^*\tilde{B}^*} = T_{\tilde{B}B}^t. \quad (\text{IV.31})$$

BEWEIS. Da die identische Abbildung id_V offensichtlich invertierbar ist, muss auch $T_{\tilde{B}B} = [\text{id}_V]_{\tilde{B}B}$ invertierbar sein und mit (IV.25) erhalten wir

$$T_{\tilde{B}B}^{-1} = [\text{id}_V]_{\tilde{B}B}^{-1} = [\text{id}_V^{-1}]_{B\tilde{B}} = [\text{id}_V]_{B\tilde{B}} = T_{B\tilde{B}},$$

also (IV.28). Gleichung (IV.29) folgt aus (IV.22), denn

$$[v]_{\tilde{B}} = [\text{id}_V(v)]_{\tilde{B}} = [\text{id}_V]_{\tilde{B}B}[v]_B = T_{\tilde{B}B}[v]_B.$$

Aus (IV.24) folgt

$$[\varphi]_{\tilde{C}\tilde{B}} = [\text{id}_W \circ \varphi \circ \text{id}_V]_{\tilde{B}\tilde{C}} = [\text{id}_W]_{\tilde{C}C}[\varphi]_{CB}[\text{id}_V]_{B\tilde{B}} = T_{\tilde{C}C}[\varphi]_{CB}T_{B\tilde{B}},$$

also (IV.30). Schließlich erhalten wir aus (IV.27) auch

$$T_{B^*\tilde{B}^*} = [\text{id}_V^*]_{B^*\tilde{B}^*} = [\text{id}_V^*]_{B^*B^*} = [\text{id}_V]_{\tilde{B}B}^t,$$

womit auch (IV.31) gezeigt wäre. \square

IV.6.22. BEMERKUNG. Sei $B = (b_1, \dots, b_n)$ eine Basis von \mathbb{K}^n und bezeichne $E = (e_1, \dots, e_n)$ die Standardbasis von \mathbb{K}^n . Dann gilt offenbar

$$T_{EB} = (b_1 | \dots | b_n),$$

und mit Korollar IV.6.21 erhalten wir erneut, vgl. Bemerkung IV.6.10,

$$[x]_B = T_{BE}[x]_E = T_{EB}^{-1}x = (b_1 | \dots | b_n)^{-1}x$$

für jedes $x \in \mathbb{K}^n$, denn $[x]_E = x$. Für die duale Basis gilt $[b_i^*]_{B^*} = e_i$, also

$$([b_j^*]_{E^*})^t = (T_{E^*B^*}[b_j^*]_{B^*})^t = (T_{BE}^t e_i)^t = e_i^t T_{BE} = e_i^t (b_1 | \dots | b_n)^{-1},$$

d.h. die Matrix die dem linearen Funktional $b_i^*: \mathbb{K}^n \rightarrow \mathbb{K}$ entspricht, $b_i^* = \psi_{a_i}$, $a_i \in M_{1 \times n}(\mathbb{K})$, ist gerade die i -te Zeile von $(b_1 | \dots | b_n)^{-1}$. Auch dies haben wir weiter oben schon festgehalten, vgl. Bemerkung IV.6.3. Ist $A \in M_{m \times n}(\mathbb{K})$ und $\varphi: \mathbb{K}^n \rightarrow \mathbb{K}^m$, $\varphi(x) = Ax$, dann gilt offenbar $[\varphi]_{EE} = A$, vgl. Bemerkung IV.6.16. Ist $C = (c_1, \dots, c_n)$ eine weitere Basis von \mathbb{K}^n so erhalten wir mittels Korollar IV.6.21

$$[\varphi]_{CB} = T_{EC}^{-1}[\varphi]_{EE}T_{EB} = (c_1 | \dots | c_n)^{-1}A(b_1 | \dots | b_n).$$

IV.6.23. BEISPIEL. Betrachte die beiden geordneten Basen

$$B = \left(\begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix}, \begin{pmatrix} -1 \\ -1 \\ 2 \end{pmatrix}, \begin{pmatrix} 1 \\ 3 \\ 3 \end{pmatrix} \right) \quad \text{und} \quad E = \left(\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right)$$

von \mathbb{R}^3 . Offensichtlich gilt dann

$$T_{EB} = \begin{pmatrix} 1 & -1 & 1 \\ 2 & -1 & 3 \\ 0 & 2 & 3 \end{pmatrix}$$

Nach (IV.28) folgt

$$T_{BE} = \begin{pmatrix} 1 & -1 & 1 \\ 2 & -1 & 3 \\ 0 & 2 & 3 \end{pmatrix}^{-1} = \begin{pmatrix} -9 & 5 & -2 \\ -6 & 3 & -1 \\ 4 & -2 & 1 \end{pmatrix}.$$

Dies erlaubt es nun die Koordinaten eines Vektors in \mathbb{R}^3 bezüglich der Basis B zu bestimmen. Betrachten wir etwa $x = \begin{pmatrix} 2 \\ 9 \\ 13 \end{pmatrix} \in \mathbb{R}^3$, dann gilt $[x]_E = \begin{pmatrix} 2 \\ 9 \\ 13 \end{pmatrix}$ also

$$[x]_B = T_{BE}[x]_E = \begin{pmatrix} -9 & 5 & -2 \\ -6 & 3 & -1 \\ 4 & -2 & 1 \end{pmatrix} \begin{pmatrix} 2 \\ 9 \\ 13 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix},$$

d.h. $\begin{pmatrix} 2 \\ 9 \\ 13 \end{pmatrix} = 1 \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix} + 2 \begin{pmatrix} -1 \\ -1 \\ 2 \end{pmatrix} + 3 \begin{pmatrix} 1 \\ 3 \\ 3 \end{pmatrix}$. Betrachte wir die lineare Abbildung

$$\varphi: \mathbb{R}^3 \rightarrow \mathbb{R}^3, \quad \varphi \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} := \begin{pmatrix} -7x_1 + 4x_2 - 2x_3 \\ 13x_2 - 6x_3 \\ 12x_2 - 5x_1 \end{pmatrix},$$

dann gilt offenbar

$$[\varphi]_{EE} = \begin{pmatrix} -7 & 4 & -2 \\ 0 & 13 & -6 \\ 0 & 12 & -5 \end{pmatrix}.$$

Mit (IV.30) können wir nun auch die Matrix von φ bezüglich B berechnen,

$$\begin{aligned} [\varphi]_{BB} &= T_{BE}[\varphi]_{EE}T_{EB} \\ &= \begin{pmatrix} -9 & 5 & -2 \\ -6 & 3 & -1 \\ 4 & -2 & 1 \end{pmatrix} \begin{pmatrix} -7 & 4 & -2 \\ 0 & 13 & -6 \\ 0 & 12 & -5 \end{pmatrix} \begin{pmatrix} 1 & -1 & 1 \\ 2 & -1 & 3 \\ 0 & 2 & 3 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix} \end{aligned}$$

Wir verstehen dadurch die Abbildung φ besser, es handelt sich um eine Spiegelung am Teilraum $\langle \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix}, \begin{pmatrix} -1 \\ -1 \\ 2 \end{pmatrix} \rangle$ längs des Teilraums $\langle \begin{pmatrix} 1 \\ 3 \\ 3 \end{pmatrix} \rangle$.

IV.6.24. BEISPIEL. Wir wollen eine lineare Abbildung $\sigma: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ bestimmen, die die Punkte auf der Geraden $\langle \begin{pmatrix} 1 \\ 2 \end{pmatrix} \rangle$ unverändert lässt und auf Punkten der Gerade $\langle \begin{pmatrix} 3 \\ 5 \end{pmatrix} \rangle$ durch Multiplikation mit -1 wirkt. Beachte, dass die beiden Vektoren $\begin{pmatrix} 1 \\ 2 \end{pmatrix}$ und $\begin{pmatrix} 3 \\ 5 \end{pmatrix}$ eine Basis von \mathbb{R}^2 bilden. Nach Proposition III.3.2 gibt es also eine eindeutig bestimmte lineare Abbildung $\sigma: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ für die $\sigma \begin{pmatrix} 1 \\ 2 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$ und $\sigma \begin{pmatrix} 3 \\ 5 \end{pmatrix} = -\begin{pmatrix} 3 \\ 5 \end{pmatrix}$ gilt. Für ihre Matrix bezüglich der Basis $B = (\begin{pmatrix} 1 \\ 2 \end{pmatrix}, \begin{pmatrix} 3 \\ 5 \end{pmatrix})$ gilt offenbar $[\sigma]_{BB} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$. Bezeichnet $E = (\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix})$ die Standardbasis dann gilt $T_{EB} = \begin{pmatrix} 1 & 3 \\ 2 & 5 \end{pmatrix}$ und daher $T_{BE} = T_{EB}^{-1} = \begin{pmatrix} -5 & 3 \\ 2 & -1 \end{pmatrix}$, siehe (IV.25). Mit (IV.24) folgt

$$[\sigma]_{EE} = T_{EB}[\sigma]_{BB}T_{BE} = \begin{pmatrix} 1 & 3 \\ 2 & 5 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} -5 & 3 \\ 2 & -1 \end{pmatrix} = \begin{pmatrix} -11 & 6 \\ -20 & 11 \end{pmatrix}$$

die gesuchte Abbildung ist daher durch $\sigma \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} := \begin{pmatrix} -11x_1 + 6x_2 \\ -20x_1 + 11x_2 \end{pmatrix}$ gegeben.

IV.6.25. BEISPIEL. Betrachte die beiden geordneten Basen

$$B = (1, z, z^2, z^3) \quad \text{und} \quad \tilde{B} = (1, z + 1, (z + 1)^2, (z + 1)^3)$$

des Vektorraums $\mathbb{K}[z]_{\leq 3}$. Dann gilt

$$T_{B\tilde{B}} = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 2 & 3 \\ 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad \text{und} \quad T_{\tilde{B}B} = T_{B\tilde{B}}^{-1} = \begin{pmatrix} 1 & -1 & 1 & -1 \\ 0 & 1 & -2 & 3 \\ 0 & 0 & 1 & -3 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

Ist etwa $p = 3 + z - 7z^2 + 9z^3$, dann folgt

$$[p]_B = \begin{pmatrix} 3 \\ 1 \\ -7 \\ 9 \end{pmatrix}, \text{ also } [p]_{\tilde{B}} = T_{\tilde{B}B}[p]_B = \begin{pmatrix} 1 & -1 & 1 & -1 \\ 0 & 1 & -2 & 3 \\ 0 & 0 & 1 & -3 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 3 \\ 1 \\ -7 \\ 9 \end{pmatrix} = \begin{pmatrix} -14 \\ 42 \\ -34 \\ 9 \end{pmatrix}.$$

Die letzte Gleichung bedeutet gerade $p = -14 + 42(z + 1) - 34(z + 1)^2 + 9(z + 1)^3$.

Wir beenden diesen Abschnitt mit folgender Anwendung:

IV.6.26. SATZ. Sei \mathbb{K} ein Körper, $x_0, \dots, x_n \in \mathbb{K}$ paarweise verschieden und $y_0, \dots, y_n \in \mathbb{K}$ beliebig. Dann existiert ein eindeutiges Polynom $p \in \mathbb{K}[z]_{\leq n}$, sodass $p(x_i) = y_i$ für alle $i = 0, \dots, n$. Insbesondere besitzt jedes nicht-triviale Polynom n -ten Grades, $0 \neq p \in \mathbb{K}[z]_{\leq n}$, höchstens n verschiedene Nullstellen in \mathbb{K} . Darüber hinaus ist die sogenannte Vandermonde-Matrix,

$$\begin{pmatrix} 1 & x_0 & x_0^2 & x_0^3 & \cdots & x_0^n \\ 1 & x_1 & x_1^2 & x_1^3 & \cdots & x_1^n \\ 1 & x_2 & x_2^2 & x_2^3 & \cdots & x_2^n \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_n & x_n^2 & x_n^3 & \cdots & x_n^n \end{pmatrix} \in M_{(n+1) \times (n+1)}(\mathbb{K}),$$

invertierbar.

BEWEIS. Für das Polynom

$$p := \sum_{i=0}^n \frac{y_i(z - x_0)(z - x_1) \cdots (z - x_{i-1})(z - x_{i+1}) \cdots (z - x_n)}{(x_i - x_0)(x_i - x_1) \cdots (x_i - x_{i-1})(x_i - x_{i+1}) \cdots (x_i - x_n)} \in \mathbb{K}[z]_{\leq n}$$

gilt offensichtlich $p(x_i) = y_i$, $i = 0, \dots, n$. Die lineare Abbildung

$$\phi: \mathbb{K}[z]_{\leq n} \rightarrow \mathbb{K}^{n+1}, \quad \phi(p) := \begin{pmatrix} p(x_0) \\ p(x_1) \\ \vdots \\ p(x_n) \end{pmatrix},$$

ist also surjektiv. Nach Korollar IV.2.11 ist ϕ daher ein Isomorphismus, denn $\dim(\mathbb{K}[z]_{\leq n}) = n + 1 = \dim(\mathbb{K}^{n+1})$. Die Injektivität von ϕ bedeutet aber gerade, dass das Polynom p durch seine Werte $p(x_0), \dots, p(x_n)$ eindeutig bestimmt ist.

Bezeichnet $E = (e_0, \dots, e_n)$ die Standardbasis von \mathbb{K}^{n+1} und $B = (1, z, z^2, \dots, z^n)$ die Basis der Monome in $\mathbb{K}[z]_{\leq n}$, dann gilt offenbar

$$[\phi]_{EB} = \begin{pmatrix} 1 & x_0 & x_0^2 & x_0^3 & \cdots & x_0^n \\ 1 & x_1 & x_1^2 & x_1^3 & \cdots & x_1^n \\ 1 & x_2 & x_2^2 & x_2^3 & \cdots & x_2^n \\ \vdots & \vdots & \vdots & \vdots & & \vdots \\ 1 & x_n & x_n^2 & x_n^3 & \cdots & x_n^n \end{pmatrix} \in M_{(n+1) \times (n+1)}(\mathbb{K}).$$

Da ϕ invertierbar ist, muss also auch die Vandermonde-Matrix invertierbar sein, siehe Satz IV.6.15. \square

Aus dem vorangehenden Satz erhalten wir sofort:

IV.6.27. KOROLLAR. *Ist \mathbb{K} ein Körper mit unendlich vielen Elementen, dann ist die Abbildung $\mathbb{K}[z] \rightarrow F(\mathbb{K}, \mathbb{K})$, die einem Polynom p die Polynomfunktion $x \mapsto p(x)$, $x \in \mathbb{K}$, zuordnet injektiv. In diesem Fall ist ein Polynom also völlig durch die entsprechende Polynomfunktion bestimmt.*

IV.6.28. BEMERKUNG. Für endliche Körper \mathbb{K} bleibt das vorangehende Resultat nicht richtig. In diesem Fall ist nämlich $F(\mathbb{K}, \mathbb{K})$ endlich-dimensional, aber $\dim(\mathbb{K}[z]) = \infty$, es kann daher keine injektive Abbildung $\mathbb{K}[z] \rightarrow F(\mathbb{K}, \mathbb{K})$ geben.

Die Tatsache, dass ein nicht-triviales Polynom n -ten Grades, $0 \neq p \in \mathbb{K}[z]_{\leq n}$, höchstens n verschiedene Nullstellen haben kann lässt sich auch auf andere Art beweisen. Ist nämlich $x_0 \in \mathbb{K}$ Nullstelle von p , dann liefert Polynomdivision ein nicht-triviales Polynom $0 \neq q \in \mathbb{K}[z]_{\leq n-1}$, sodass $p = (z - x_0)q$. Ist $x_1 \in \mathbb{K}$ eine weitere Nullstelle von p und $x_0 \neq x_1$, dann muss x_1 eine Nullstelle von q sein und wir können einen weiteren Linearfaktor abspalten. Da der Grad des Polynoms q strikt kleiner als der Grad von p ist, folgt daraus, dass p höchstens n verschiedene Nullstellen haben kann.

V. Determinanten

Wir werden in diesem Kapitel jeder quadratischen Matrix $A \in M_{n \times n}(\mathbb{K})$ ihre Determinante, $\det(A) \in \mathbb{K}$, zuordnen. Die Determinante ist multiplikativ, es gilt

$$\det(AB) = \det(A) \det(B),$$

für je zwei $(n \times n)$ -Matrizen A und B . Auch ist eine Matrix A genau dann invertierbar, wenn $\det(A) \neq 0$ gilt. Es ist keineswegs offensichtlich, dass eine Abbildung mit diesen Eigenschaften überhaupt existiert.

In Abschnitt V.1 werden wir eine (rekursive) Konstruktion der Determinante besprechen, ihre wichtigsten Eigenschaften herleiten, und auch einen effektiven Algorithmus zur Berechnung kennen lernen. Dort werden wir auch die geometrische Bedeutung reeller Determinante klären, sie stimmt mit dem orientierten Volumen des von den Spalten der Matrix aufgespannten Parallelepipeds überein. Aus diesem Grund tritt die Determinante in der Transformationsformel für Mehrfachintegrale auf.

Mit Hilfe von Determinanten kann die Inverse einer invertierbaren Matrix A explizit angegeben werden, und die Cramer'sche Regel liefert eine explizite Lösungsformel für das lineare Gleichungssystem $Ax = y$, siehe Abschnitt V.2.

In Abschnitt V.3 werden wir eine nach Leibniz benannte explizite Formel für die Determinante einer Matrix angeben, und so eine weitere, unabhängige Konstruktion der Determinantenfunktion kennen lernen. Diese Formel besteht aus sehr vielen Termen und ist daher bei der Berechnung großer Determinanten nicht hilfreich. Für (3×3) -Matrizen liefert sie die Regel von Sarrus:

$$\det \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} = a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} \\ - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33} - a_{13}a_{22}a_{31}.$$

In Abschnitt V.4 werden wir den Begriff der Determinante auf lineare Abbildung $\varphi: V \rightarrow V$ ausdehnen, wobei V einen endlich-dimensionalen Vektorraum bezeichnet. Im letzten Abschnitt V.5 werden wir mit Hilfe von Determinanten den Begriff der Orientierungen eines endlich-dimensionalen reellen Vektorraums präzise fassen und kurz auf die wichtigsten Eigenschaften eingehen.

Vieles in diesem Kapitel findet sich etwa [3].

V.1. Determinantenfunktionen. Eine Funktion

$$\delta: M_{n \times n}(\mathbb{K}) \rightarrow \mathbb{K}, \quad A \mapsto \delta(A),$$

wird *linear in der k -ten Spalte* genannt, falls für alle $a_i \in \mathbb{K}^n$ die Abbildung

$$\mathbb{K}^n \rightarrow \mathbb{K}, \quad x \mapsto \delta(a_1 | \cdots | a_{k-1} | x | a_{k+1} | \cdots | a_n)$$

linear ist, d.h. wenn für alle $a_i, x, y \in \mathbb{K}^n$ und $\lambda \in \mathbb{K}$ gilt:

$$\begin{aligned}\delta(\cdots |a_{k-1}|x+y|a_{k+1}|\cdots) &= \delta(\cdots |a_{k-1}|x|a_{k+1}|\cdots) + \delta(\cdots |a_{k-1}|y|a_{k+1}|\cdots) \\ \delta(\cdots |a_{k-1}|\lambda x|a_{k+1}|\cdots) &= \lambda \delta(\cdots |a_{k-1}|x|a_{k+1}|\cdots)\end{aligned}$$

Gilt dies für jedes $k = 1, \dots, n$, dann wird δ *multilinear* genannt, wir sagen auch $\delta(A)$ ist linear in jeder Spalte von A . Offensichtlich bilden die multilinearen Abbildungen $M_{n \times n}(\mathbb{K}) \rightarrow \mathbb{K}$ einen Teilraum des Vektorraums aller \mathbb{K} -wertigen Funktion auf $M_{n \times n}(\mathbb{K})$, siehe Aufgabe 1.

V.1.1. LEMMA. Sei \mathbb{K} ein Körper, $n \in \mathbb{N}$ und $\delta: M_{n \times n}(\mathbb{K}) \rightarrow \mathbb{K}$ eine Funktion, die folgende beiden Eigenschaften hat:

- (a) δ ist multilinear, d.h. $\delta(A)$ ist linear in jeder Spalte von A .
- (b) Sind zwei benachbarte Spalten von A gleich, so gilt $\delta(A) = 0$.

Dann besitzt δ auch folgende Eigenschaften:

- (c) Bei Vertauschung zweier Spalten von A wechselt $\delta(A)$ das Vorzeichen.
- (d) Sind zwei beliebige Spalten von A gleich, so gilt $\delta(A) = 0$.
- (e) Addieren wir ein Vielfaches einer Spalte von A zu einer anderen Spalte, dann bleibt $\delta(A)$ unverändert.
- (f) Ist $\text{rank}(A) < n$, dann gilt $\delta(A) = 0$.
- (g) Für jedes $j \in \{1, \dots, n\}$ und jede Matrix A gilt

$$\delta(A) = \sum_{i=1}^n \delta(A_{(ij)}).$$

Dabei bezeichnet $A_{(ij)}$ jene $(n \times n)$ -Matrix, die wir aus A erhalten wenn wir jeden Eintrag in der i -ten Zeile und j -ten Spalte, außer A_{ij} , durch 0 ersetzen.⁴

- (h) Bezeichnet $\delta': M_{n \times n}(\mathbb{K}) \rightarrow \mathbb{K}$ eine weitere Funktion, die die beide die Eigenschaften (a) und (b) hat, und ist $\delta(I_n) = \delta'(I_n)$, dann gilt schon $\delta(A) = \delta'(A)$, für alle A .
- (i) Es gilt $\delta(BA)\delta(I_n) = \delta(B)\delta(A)$, für je zwei Matrizen A und B .
- (j) Für jedes $i \in \{1, \dots, n\}$ und jede Matrix A gilt

$$\delta(A) = \sum_{j=1}^n \delta(A_{(ij)}).$$

⁴D.h. $A_{(ij)} = \begin{pmatrix} & & & 0 & & & \\ & & & \vdots & & & \\ & & & 0 & & & \\ 0 & \cdots & 0 & A_{ij} & 0 & \cdots & 0 \\ & & & 0 & & & \\ & & & \vdots & & & \\ & & & 0 & & & \end{pmatrix}$, wobei A_{ij} den Eintrag von A in der i -ten

Zeile und j -ten Spalte bezeichnet. Die mit A symbolisierten Blöcke stehen jeweils für den entsprechenden Teil von A , diese bleiben unverändert.

- (k) $\delta(A)$ ist linear in jeder Zeile von A .
 (l) $\delta(A^t) = \delta(A)$, für jede Matrix A .
 (m) Sind zwei beliebige Zeilen von A gleich, so gilt $\delta(A) = 0$.
 (n) Bei Vertauschung zweier Zeilen von A wechselt $\delta(A)$ das Vorzeichen.
 (o) Addieren wir ein Vielfaches einer Zeile von A zu einer anderen Zeile, dann bleibt $\delta(A)$ unverändert.

BEWEIS. Wir beobachten zunächst, dass $\delta(A)$ bei Vertauschung zweier benachbarter Spalten das Vorzeichen wechselt, d.h. es gilt

$$\delta(\cdots |x|y|\cdots) = -\delta(\cdots |y|x|\cdots),$$

denn aus (a) und (b) folgt sofort:

$$\begin{aligned} 0 &= \delta(\cdots |x+y|x+y|\cdots) \\ &= \delta(\cdots |x|x|\cdots) + \delta(\cdots |x|y|\cdots) + \delta(\cdots |y|x|\cdots) + \delta(\cdots |y|y|\cdots) \\ &= \delta(\cdots |x|y|\cdots) + \delta(\cdots |y|x|\cdots). \end{aligned}$$

Sind nun $1 \leq i < j \leq n$, dann folgt durch sukzessives Vertauschen von Spalten:

$$\begin{aligned} \delta(\cdots \overbrace{|x|\cdots|y|}^{\text{Sp. } i \text{ bis } j} \cdots) &= (-1)^{j-i-1} \delta(\cdots |x|y|\cdots) \\ &= -(-1)^{j-i-1} \delta(\cdots |y|x|\cdots) = -\delta(\cdots \underbrace{|y|\cdots|x|}_{\text{Sp. } i \text{ bis } j} \cdots) \end{aligned}$$

Dies zeigt (c). Zusammen mit (b) erhalten wir daraus auch (d), denn

$$\delta(\cdots |x|\cdots|x|\cdots) = \pm \delta(\cdots |x|x|\cdots) = 0.$$

Unter Verwendung von (a) folgt weiters

$$\begin{aligned} \delta(\cdots |x|\cdots|y+\lambda x|\cdots) &= \delta(\cdots |x|\cdots|y|\cdots) + \lambda \delta(\cdots |x|\cdots|x|\cdots) \\ &= \delta(\cdots |x|\cdots|y|\cdots), \end{aligned}$$

für jedes $\lambda \in \mathbb{K}$. Somit ist auch (e) gezeigt.

Ad (f): Gilt $\text{rank}(A) < n$, dann sind die Spalten von $A = (a_1 | \cdots | a_n)$ linear abhängig, wenigstens eine lässt sich daher als Linearkombination der anderen schreiben. Durch Vertauschen zweier geeigneter Spalten dürfen wir o.B.d.A. $a_1 = \sum_{i=2}^n \lambda_i a_i$ annehmen, siehe (c). Mit (a) und (d) folgt dann

$$\delta(A) = \delta\left(\sum_{i=2}^n \lambda_i a_i \mid a_2 \mid a_3 \mid \cdots \mid a_n\right) = \sum_{i=2}^n \lambda_i \delta(a_i | a_2 | a_3 | \cdots | a_n) = 0.$$

Ad (g): Für die j -te Spalte von A gilt $a_j = \sum_{i=1}^n A_{ij} e_i$, wobei A_{ij} den Eintrag in der i -ten Zeile und j -ten Spalte von A bezeichnet. Mit (a) und (e) folgt

$$\delta(A) = \sum_{i=1}^n \delta(a_1 | \cdots | a_{j-1} | A_{ij} e_i | a_{j+1} | \cdots | a_n) = \sum_{i=1}^n \delta(A_{ij}),$$

denn es gilt

$$\delta(a_1 | \cdots | a_{j-1} | A_{ij} e_i | a_{j+1} | \cdots | a_n) = \delta(A_{(ij)}). \quad (\text{V.1})$$

Für $A_{ij} = 0$ ist die Gleichung (V.1) offensichtlich, denn in diesem Fall haben beide Matrizen eine verschwindende j -te Spalte, nach (a) verschwinden daher beide Seiten in (V.1). Ist $A_{ij} \neq 0$, dann können wir durch Addition geeigneter Vielfacher der j -ten Spalte von $(a_1 | \cdots | a_{j-1} | A_{ij} e_i | a_{j+1} | \cdots | a_n)$ zu den anderen Spalten diese Matrix zu $A_{(ij)}$ umformen, und dabei bleibt δ unverändert, vgl. (e).

Ad (h): Beachte zunächst, dass auch die Funktion

$$\delta'': M_{n \times n}(\mathbb{K}) \rightarrow \mathbb{K}, \quad \delta''(A) := \delta(A) - \delta'(A),$$

die beiden Eigenschaften (a) und (b) besitzt. Wie wir oben gezeigt haben, hat sie daher auch Eigenschaft (c) und (d). Nach Voraussetzung gilt weiters $\delta''(I_n) = 0$. Sei nun A eine $(n \times n)$ -Matrix und bezeichne A_{ij} den Eintrag in der i -ten Zeile und j -ten Spalte. Aus (a) folgt

$$\begin{aligned} \delta''(A) &= \delta'' \left(\sum_{i_1=1}^n A_{i_1,1} e_{i_1} \mid \cdots \mid \sum_{i_n=1}^n A_{i_n,n} e_{i_n} \right) \\ &= \sum_{i_1=1}^n \cdots \sum_{i_n=1}^n A_{i_1,1} \cdots A_{i_n,n} \delta''(e_{i_1} | \cdots | e_{i_n}), \end{aligned}$$

es genügt daher $\delta''(e_{i_1} | \cdots | e_{i_n}) = 0$ zu zeigen, für alle $i_1, \dots, i_n \in \{1, \dots, n\}$. Nach (d) genügt es dies für paarweise verschiedene i_1, \dots, i_n zu verifizieren. In diesem Fall folgt durch sukzessives Vertauschen mittels (c)

$$\delta''(e_{i_1} | \dots | e_{i_n}) = \pm \delta''(e_1 | \dots | e_n) = \pm \delta''(I_n) = 0.$$

Dies zeigt $\delta''(A) = 0$, für jedes A , und daher $\delta(A) = \delta'(A)$.

Ad (i): Sei B fix. Beachte, dass die Funktionen

$$\delta': M_{n \times n}(\mathbb{K}) \rightarrow \mathbb{K}, \quad \delta'(A) := \delta(BA)\delta(I_n)$$

und

$$\delta'': M_{n \times n}(\mathbb{K}) \rightarrow \mathbb{K}, \quad \delta''(A) := \delta(B)\delta(A)$$

beide die Eigenschaften (a) und (b) haben. Bezeichnen nämlich $A = (a_1 | \cdots | a_n)$ die Spalten von A dann gilt $BA = (Ba_1 | \cdots | Ba_n)$, also hängt die k -te Spalte von BA linear von der k -ten Spalte von A ab, und, falls die k -te Spalte von B mit der l -ten Spalte übereinstimmt, dann bleibt dies auch für BA richtig. Da offensichtlich auch $\delta'(I_n) = \delta''(I_n)$ gilt, erhalten wir aus (h) nun $\delta'(A) = \delta''(A)$, d.h. $\delta(BA)\delta(I_n) = \delta(B)\delta(A)$.

Ad (j): Beachte, dass die k -te Spalte von $A_{(ij)}$ linear von der k -ten Spalte von A abhängt. Daher stellt auch

$$\delta': M_{n \times n}(\mathbb{K}) \rightarrow \mathbb{K}, \quad \delta'(A) := \sum_{j=1}^n \delta(A_{(ij)})$$

eine Funktion dar, die linear in den Spalten von A ist. Stimmen die Spalten mit Nummern k und l von A überein, $1 \leq k < l \leq n$, dann folgt

$$\delta'(A) = \delta(A_{(ik)}) + \delta(A_{(il)}) = 0,$$

denn für $j \neq k, l$ hat auch $A_{(ij)}$ zwei gleiche Spalten, also $\delta(A_{(ij)}) = 0$ nach (d), und die Matrix $A_{(il)}$ entsteht aus $A_{(ik)}$ durch Vertauschen der Spalten k und l , also $\delta(A_{(ik)}) = -\delta(A_{(il)})$ wegen (c). Schließlich gilt auch $\delta'(I_n) = \delta(I_n)$, denn $(I_n)_{(ii)} = I_n$ und für $j \neq i$ ist $\delta((I_n)_{(ij)}) = 0$, da die j -te Spalte von $(I_n)_{(ij)}$ verschwindet. Aus (h) erhalten wir daher $\delta(A) = \delta'(A)$, d.h. $\delta(A) = \sum_{j=1}^n \delta(A_{(ij)})$.

Ad (k): Beachte zunächst, dass die i -te Spalte von $A_{(ij)}$ linear von der i -ten Zeile von A abhängt. Folglich ist $\delta(A_{(ij)})$ linear in der i -ten Zeile von A . Nach (j) ist daher auch $\delta(A)$ linear in der i -ten Zeile von A . Da i beliebig war, folgt die Behauptung.

Ad (l): Nach (k) ist die Funktion

$$\delta': M_{n \times n}(\mathbb{K}) \rightarrow \mathbb{K}, \quad \delta'(A) := \delta(A^t),$$

linear in jeder Spalte. Ist $\text{rank}(A) < n$, dann auch $\text{rank}(A^t) < n$, also $\delta'(A) = \delta(A^t) = 0$ nach (f). Insbesondere verschwindet $\delta'(A)$, falls zwei Spalten von A übereinstimmen. Mit (h) folgt somit $\delta'(A) = \delta(A)$, denn offensichtlich gilt auch $\delta'(I_n) = \delta(I_n)$.

Die verbleibenden Behauptungen (m), (n) und (o) folgen sofort aus (c), (d) und (e) mittels (m). \square

V.1.2. BEMERKUNG. Nach Lemma V.1.1 sind für eine multilineare Abbildung $\delta: M_{n \times n}(\mathbb{K}) \rightarrow \mathbb{K}$ folgende Aussagen äquivalent:

- (a) Stimmen zwei benachbarte Spalten von A überein, so gilt $\delta(A) = 0$.
- (b) Stimmen zwei Spalten von A überein, so gilt $\delta(A) = 0$.
- (c) Sind die Spalten von A linear abhängig, so gilt $\delta(A) = 0$.

Eine multilineare Abbildung $\delta: M_{n \times n}(\mathbb{K}) \rightarrow \mathbb{K}$, die diese Eigenschaften besitzt wird *alternierende multilineare Abbildung* genannt. Offensichtlich bilden die alternierenden multilinearen Abbildungen einen Teilraum des Vektorraums aller multilinearen Abbildungen $M_{n \times n}(\mathbb{K}) \rightarrow \mathbb{K}$, vgl. Aufgabe 1. In Satz V.1.3 unten werden wir eine alternierende multilineare Abbildung $\det: M_{n \times n}(\mathbb{K}) \rightarrow \mathbb{K}$ konstruieren, für die $\det(I_n) = 1$ gilt. Nach Lemma V.1.1(h) ist daher jede alternierende multilineare Abbildung $\delta: M_{n \times n}(\mathbb{K}) \rightarrow \mathbb{K}$ von der Form $\delta(A) = \delta(I_n) \det(A)$. Der Vektorraum der alternierenden multilinearen Abbildungen $M_{n \times n}(\mathbb{K}) \rightarrow \mathbb{K}$ hat daher Dimension 1 und \det bildet eine Basis. Gilt $2 \neq 0 \in \mathbb{K}$, dann sind die Eigenschaften (a) bis (c) oben auch zu folgender Bedingung äquivalent:

- (d) Bei Vertauschung zweier Spalten von A wechselt $\delta(A)$ das Vorzeichen.

Für allgemeine Körper bleibt dies jedoch nicht richtig, siehe Aufgabe 2.

V.1.3. SATZ. Sei \mathbb{K} ein Körper und $n \in \mathbb{N}$. Dann existiert eine eindeutige Funktion $\det: M_{n \times n}(\mathbb{K}) \rightarrow \mathbb{K}$, die folgende drei Eigenschaften besitzt:

- (a) \det ist multilinear, d.h. $\det(A)$ ist linear in jeder Spalte von A .
 (b) Sind zwei benachbarte Spalten von A gleich, dann gilt $\det(A) = 0$.
 (c) $\det(I_n) = 1$, wobei I_n die Einheitsmatrix bezeichnet.

Diese Funktion hat darüber hinaus folgende Eigenschaften:

- (d) Bei Vertauschung zweier Spalten von A wechselt $\det(A)$ das Vorzeichen.
 (e) Addieren wir ein Vielfaches einer Spalte von A zu einer anderen Spalte, so bleibt $\det(A)$ unverändert.
 (f) $\det(A)$ ist linear in jeder Zeile von A .
 (g) Bei Vertauschung zweier Zeilen von A wechselt $\det(A)$ das Vorzeichen.
 (h) Addieren wir ein Vielfaches einer Zeile von A zu einer anderen Zeile, so bleibt $\det(A)$ unverändert.
 (i) Für je zwei Matrizen A und B gilt

$$\det(AB) = \det(A) \det(B).$$

- (j) A ist genau dann invertierbar, wenn $\det(A) \neq 0$. In diesem Fall gilt

$$\det(A^{-1}) = \det(A)^{-1}.$$

- (k) Für jede Matrix A ist

$$\det(A^t) = \det(A).$$

- (l) Für jedes $i \in \{1, \dots, n\}$ und jede Matrix A gilt

$$\det(A) = \sum_{j=1}^n (-1)^{i+j} A_{ij} \det(\hat{A}_{(ij)}).$$

Diese Formel wird Entwicklung nach der i -ten Zeile genannt. Dabei bezeichnet A_{ij} den Eintrag der Matrix A in der i -ten Zeile und j -ten Spalte, und $\hat{A}_{(ij)}$ bezeichnet jene $(n-1) \times (n-1)$ -Matrix, die wir durch Streichen der i -ten Zeile und j -ten Spalte aus A gewinnen.

- (m) Für jedes $j \in \{1, \dots, n\}$ und jede Matrix A gilt

$$\det(A) = \sum_{i=1}^n (-1)^{i+j} A_{ij} \det(\hat{A}_{(ij)}).$$

Diese Formel wird Entwicklung nach der j -ten Spalte genannt.

- (n) Für Dreiecksmatrizen haben wir

$$\det \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ 0 & a_{22} & \cdots & a_{2n} \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & a_{nn} \end{pmatrix} = a_{11} a_{22} \cdots a_{nn} = \det \begin{pmatrix} a_{11} & 0 & \cdots & 0 \\ a_{21} & a_{22} & \ddots & \vdots \\ \vdots & \vdots & \ddots & 0 \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}$$

BEWEIS. Wir zeigen zunächst mittels Induktion nach n , dass tatsächlich eine Funktion $\det: M_{n \times n}(\mathbb{K}) \rightarrow \mathbb{K}$ mit den Eigenschaften (a), (b) und (c) existiert. Der Fall $n = 1$ ist trivial, die Abbildung $\det: M_{1 \times 1}(\mathbb{K}) \rightarrow \mathbb{K}$, $\det(a) := a$, hat offensichtlich alle gewünschten Eigenschaften. Für den Induktionsschritt sei nun

$\det: M_{(n-1) \times (n-1)}(\mathbb{K}) \rightarrow \mathbb{K}$ eine Abbildung, sodass (a), (b) und (c) gelten. Weiters fixieren wir ein $i \in \{1, \dots, n\}$. Für $A \in M_{n \times n}(\mathbb{K})$ definieren wir nun

$$\det(A) := \sum_{j=1}^n (-1)^{i+j} A_{ij} \det(\hat{A}_{(ij)}). \quad (\text{V.2})$$

Dabei bezeichnet A_{ij} den Eintrag der Matrix A in der i -ten Zeile und j -ten Spalte, und $\hat{A}_{(ij)}$ bezeichnet jene $(n-1) \times (n-1)$ -Matrix, die wir durch Streichen der i -ten Zeile und j -ten Spalte aus A erhalten. Jedenfalls ist dadurch eine Abbildung $\det: M_{n \times n}(\mathbb{K}) \rightarrow \mathbb{K}$ definiert. Wir verifizieren nun, dass diese die Eigenschaften (a), (b) und (c) hat.

Ad (a): Seien $j, k \in \{1, \dots, n\}$. Es genügt zu zeigen, dass $A_{ij} \det(\hat{A}_{(ij)})$ linear in der k -ten Spalte von A ist, vgl. (V.2). Wir unterscheiden zwei Fälle. Für $k = j$ hängt $\det(\hat{A}_{(ij)})$ nicht von der k -ten Spalte ab, denn diese wurde ja gestrichen, und A_{ij} ist offensichtlich linear in der k -ten Spalte. Ist $j \neq k$, so ist A_{ij} unabhängig von der k -ten Spalte und nach Induktionsvoraussetzung hängt $\det(\hat{A}_{(ij)})$ linear von der k -ten Spalte von A ab. In beiden Fällen folgt, dass $A_{ij} \det(\hat{A}_{(ij)})$ linear in der k -ten Spalte von A ist.

Ad (b): Sei $A \in M_{n \times n}(\mathbb{K})$ eine Matrix deren k -te Spalte mit der $(k+1)$ -ten Spalte übereinstimmt, $1 \leq k < n$. Ist nun $j \neq k$ und $j \neq k+1$, dann hat auch $\hat{A}_{(ij)}$ zwei gleiche benachbarte Spalten, nach Induktionsvoraussetzung gilt daher $\det(\hat{A}_{(ij)}) = 0$ für diese j . In der Summe (V.2) bleiben daher nur zwei Terme übrig,

$$\det(A) = (-1)^{i+k} A_{ik} \det(\hat{A}_{(ik)}) + (-1)^{i+k+1} A_{i,k+1} \det(\hat{A}_{(i,k+1)}).$$

Weiters ist $\hat{A}_{(ik)} = \hat{A}_{(i,k+1)}$, die beiden Terme auf der rechten Seite stimmen also bis auf ihr Vorzeichen überein, und es folgt $\det(A) = 0$.

Ad (c): Der einzige nicht verschwindende Eintrag der i -ten Zeile der Einheitsmatrix I_n ist $(I_n)_{ii} = 1$. Weiters gilt $(\hat{I}_n)_{(ii)} = I_{n-1}$ und nach Induktionsvoraussetzung daher $\det((\hat{I}_n)_{(ii)}) = \det(I_{n-1}) = 1$. Aus (V.2) erhalten wir daher

$$\det(I_n) = \sum_{j=1}^n (-1)^{i+j} (I_n)_{ij} \det((\hat{I}_n)_{(ij)}) = (-1)^{i+i} (I_n)_{ii} \det((\hat{I}_n)_{(ii)}) = 1.$$

Damit ist der Induktionsschritt gezeigt, also existiert für jedes $n \in \mathbb{N}$ eine Funktion $\det: M_{n \times n}(\mathbb{K}) \rightarrow \mathbb{K}$, die die Eigenschaften (a), (b) und (c) besitzt. Die Eindeutigkeit dieser Funktion folgt aus Lemma V.1.1(h). Aus (V.2) erhalten wir auch sofort (l).

Die Behauptungen (d), (e), (f), (g), (h), (i) und (k) folgen sofort aus den entsprechenden Aussagen in Lemma V.1.1.

Ad (j): Ist A invertierbar, so erhalten wir mit (c) und (i)

$$1 = \det(I_n) = \det(AA^{-1}) = \det(A) \det(A^{-1}).$$

Es folgt $\det(A) \neq 0$ und $\det(A^{-1}) = \det(A)^{-1}$. Ist A nicht invertierbar, dann gilt $\text{rank}(A) < n$ und daher $\det(A) = 0$ nach Lemma V.1.1(f).

Ad (m): Dies folgt aus Lemma V.1.1(g), denn mit (l) erhalten wir $\det(A_{(ij)}) = (-1)^{i+j} A_{ij} \det(\hat{A}_{(ij)})$. Alternativ lässt sich (m) auch aus (l) mittels (k) herleiten.

Ad (n): Durch Entwicklung nach der ersten Zeile, siehe (l), erhalten wir

$$\det \begin{pmatrix} a_{11} & 0 & \cdots & 0 \\ a_{21} & a_{22} & \ddots & \vdots \\ \vdots & \vdots & \ddots & 0 \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix} = a_{11} \det \begin{pmatrix} a_{22} & 0 & \cdots & 0 \\ a_{32} & a_{33} & \ddots & \vdots \\ \vdots & \vdots & \ddots & 0 \\ a_{n2} & a_{n3} & \cdots & a_{nn} \end{pmatrix}.$$

Mittels Induktion nach n folgt die zweite Gleichung in (n). Die erste lässt sich analog beweisen oder mit Hilfe von (k) aus der eben bewiesenen ableiten. \square

V.1.4. DEFINITION (Determinante einer Matrix). Die Abbildung

$$\det: M_{n \times n}(\mathbb{K}) \rightarrow \mathbb{K}, \quad A \mapsto \det(A),$$

aus Satz V.1.3 wird die *Determinantenfunktion* genannt, und $\det(A)$ als *Determinante* der Matrix A bezeichnet. Eine andere gängige Notation ist:

$$\begin{vmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{vmatrix} := \det \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix}$$

V.1.5. BEISPIEL. Für 2×2 -Matrizen erhalten wir durch Entwicklung nach der ersten Zeile, siehe Satz V.1.3(l),

$$\det \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = a_{11}a_{22} - a_{12}a_{21}.$$

V.1.6. BEISPIEL (Regel von Sarrus). Für 3×3 -Matrizen gilt:

$$\det \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} = a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} \\ - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33} - a_{13}a_{22}a_{31}$$

Auch dies folgt durch Entwicklung nach der ersten Zeile, siehe Satz V.1.3(l), zusammen mit der eben hergeleiteten Formel für die Determinante einer 2×2 -Matrix:

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = a_{11} \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - a_{12} \begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} + a_{13} \begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix}$$

V.1.7. BEMERKUNG. Analog lassen sich rekursiv Formeln für größere Matrizen herleiten, etwa gilt:

$$\begin{aligned} \det \begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{pmatrix} &= a_{11} \begin{vmatrix} a_{22} & a_{23} & a_{24} \\ a_{32} & a_{33} & a_{34} \\ a_{42} & a_{43} & a_{44} \end{vmatrix} - a_{12} \begin{vmatrix} a_{21} & a_{23} & a_{24} \\ a_{31} & a_{33} & a_{34} \\ a_{41} & a_{43} & a_{44} \end{vmatrix} \\ &\quad + a_{13} \begin{vmatrix} a_{21} & a_{22} & a_{24} \\ a_{31} & a_{32} & a_{34} \\ a_{41} & a_{42} & a_{44} \end{vmatrix} - a_{14} \begin{vmatrix} a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \\ a_{41} & a_{42} & a_{43} \end{vmatrix} \\ &= a_{11}a_{22}a_{33}a_{44} + a_{11}a_{23}a_{34}a_{42} + a_{11}a_{24}a_{32}a_{43} \\ &\quad - a_{11}a_{22}a_{34}a_{43} - a_{11}a_{23}a_{32}a_{44} - a_{11}a_{24}a_{33}a_{42} \\ &\quad \pm 18 \text{ weitere ähnliche Terme.} \end{aligned}$$

Die Anzahl der Terme wächst rapide, für eine $(n \times n)$ -Matrix sind es $n!$ Summanden. In Abschnitt V.3 werden wir einen geschlossenen Ausdruck dafür angeben, siehe Satz V.3.6. Diese Formeln sind für die Berechnung der Determinanten nur von begrenzter Nützlichkeit.

V.1.8. BEMERKUNG (Volumen und Determinanten). Der orientierte Flächeninhalt des von zwei Vektoren $a_1, a_2 \in \mathbb{R}^2$ aufgespannten Parallelogramms ist $\det(a_1|a_2)$. Um dies einzusehen, bezeichne $\delta(a_1|a_2)$ den orientierten Flächeninhalt dieses Parallelogramms. Offensichtlich gilt $\delta(I_2) = 1$, denn das Einheitsquadrat hat Fläche 1. Auch gilt $\delta(a|a) = 0$, denn in diesem Fall degeneriert das Parallelogramm zu einer Linie ohne Flächeninhalt. Eine einfache elementargeometrische Überlegung zeigt, dass $\delta(a_1|a_2)$ linear in beiden Spalten ist. Nach Satz V.1.3 muss daher $\delta(a_1|a_2) = \det(a_1|a_2)$ gelten, die Determinante $\det(a_1|a_2)$ berechnet also den orientierten Flächeninhalt des von a_1 und a_2 aufgespannten Parallelogramms. Völlig analog lässt sich zeigen, dass das orientierte Volumen des von drei Vektoren $b_1, b_2, b_3 \in \mathbb{R}^3$ aufgespannten Parallelepipeds durch $\det(b_1|b_2|b_3)$ gegeben ist.

V.1.9. BEMERKUNG. Mit Satz V.1.3(j) erhalten wir folgende Beschreibung der allgemeinen linearen Gruppe durch eine (nicht lineare) Ungleichung:

$$\mathrm{GL}_n(\mathbb{K}) = \{A \in M_{n \times n}(\mathbb{K}) : \det(A) \neq 0\}.$$

Die Determinante schränkt sich zu einem surjektiven Gruppenhomomorphismus,

$$\det: \mathrm{GL}_n(\mathbb{K}) \rightarrow \mathbb{K} \setminus \{0\},$$

ein, wobei $\mathbb{K} \setminus \{0\}$ mit der Multiplikation des Körpers als (abelsche) Gruppe aufgefasst wird. Der Kern dieses Homomorphismus,

$$\mathrm{SL}_n(\mathbb{K}) := \{A \in M_{n \times n}(\mathbb{K}) : \det(A) = 1\},$$

bildet daher eine Untergruppe von $\mathrm{GL}_n(\mathbb{K})$, sie wird die *spezielle lineare Gruppe* genannt und ist durch eine (nicht lineare) Gleichung beschrieben.

Soll die Determinante einer Matrix berechnet werden ist es zweckmäßig diese durch Zeilen und oder Spaltenumformungen auf Dreiecksgestalt zu bringen und dann die Formel aus Satz V.1.3(n) zu verwenden.

V.1.10. BEISPIEL. Wir wollen die Determinante der Matrix

$$A = \begin{pmatrix} 1 & 5 & 8 & 10 \\ -1 & -5 & 1 & 6 \\ 2 & 12 & 22 & 29 \\ 3 & 17 & 30 & 43 \end{pmatrix}$$

bestimmen. Mit den Rechenregeln in Satz V.1.3 folgt

$$\begin{vmatrix} 1 & 5 & 8 & 10 \\ -1 & -5 & 1 & 6 \\ 2 & 12 & 22 & 29 \\ 3 & 17 & 30 & 43 \end{vmatrix} = \begin{vmatrix} 1 & 5 & 8 & 10 \\ 0 & 0 & 9 & 16 \\ 0 & 2 & 6 & 9 \\ 0 & 2 & 6 & 13 \end{vmatrix} = - \begin{vmatrix} 1 & 5 & 8 & 10 \\ 0 & 2 & 6 & 9 \\ 0 & 0 & 9 & 16 \\ 0 & 2 & 6 & 13 \end{vmatrix} = - \begin{vmatrix} 1 & 5 & 8 & 10 \\ 0 & 2 & 6 & 9 \\ 0 & 0 & 9 & 16 \\ 0 & 0 & 0 & 4 \end{vmatrix} = -1 \cdot 2 \cdot 9 \cdot 4 = -72,$$

d.h. $\det(A) = -72$. Beachte, dass diese Vorgehensweise wesentlich effektiver ist, als Einsetzen in die Formel aus Bemerkung V.1.7.

V.1.11. BEISPIEL. Wir wollen die Determinante der Matrix

$$A = \begin{pmatrix} 1 & 7 & 1 & 2 & 1 \\ 1 & 9 & 3 & 5 & 3 \\ -1 & -5 & 4 & 7 & 4 \\ -2 & -14 & 1 & 6 & 5 \\ 1 & 5 & -4 & -3 & 5 \end{pmatrix}$$

bestimmen. Wie zuvor folgt mit Satz V.1.3

$$\begin{vmatrix} 1 & 7 & 1 & 2 & 1 \\ 1 & 9 & 3 & 5 & 3 \\ -1 & -5 & 4 & 7 & 4 \\ -2 & -14 & 1 & 6 & 5 \\ 1 & 5 & -4 & -3 & 5 \end{vmatrix} = \begin{vmatrix} 1 & 7 & 1 & 2 & 1 \\ 0 & 2 & 2 & 3 & 2 \\ 0 & 2 & 5 & 9 & 5 \\ 0 & 0 & 3 & 10 & 7 \\ 0 & -2 & -5 & -5 & 4 \end{vmatrix} = \begin{vmatrix} 1 & 7 & 1 & 2 & 1 \\ 0 & 2 & 2 & 3 & 2 \\ 0 & 0 & 3 & 6 & 3 \\ 0 & 0 & 3 & 10 & 7 \\ 0 & 0 & -3 & -2 & 6 \end{vmatrix} \\ = \begin{vmatrix} 1 & 7 & 1 & 2 & 1 \\ 0 & 2 & 2 & 3 & 2 \\ 0 & 0 & 3 & 6 & 3 \\ 0 & 0 & 4 & 4 & 4 \\ 0 & 0 & 4 & 9 & 5 \end{vmatrix} = \begin{vmatrix} 1 & 7 & 1 & 2 & 1 \\ 0 & 2 & 2 & 3 & 2 \\ 0 & 0 & 3 & 6 & 3 \\ 0 & 0 & 4 & 4 & 4 \\ 0 & 0 & 0 & 5 & 5 \end{vmatrix} = 1 \cdot 2 \cdot 3 \cdot 4 \cdot 5 = 120,$$

also $\det(A) = 120$.

V.1.12. BEISPIEL. Wir wollen die Determinante der komplexen Matrix

$$A = \begin{pmatrix} \mathbf{i} & 7 + \mathbf{i} & 3 - 2\mathbf{i} \\ 2\mathbf{i} & 15 + 3\mathbf{i} & 8 - \mathbf{i} \\ 1 & 2 - 6\mathbf{i} & 1 - \mathbf{i} \end{pmatrix}$$

bestimmen. Mit Satz V.1.3 erhalten wir

$$\begin{vmatrix} \mathbf{i} & 7+\mathbf{i} & 3-2\mathbf{i} \\ 2\mathbf{i} & 15+3\mathbf{i} & 8-\mathbf{i} \\ 1 & 2-6\mathbf{i} & 1-\mathbf{i} \end{vmatrix} = \begin{vmatrix} \mathbf{i} & 7+\mathbf{i} & 3-2\mathbf{i} \\ 0 & 1+\mathbf{i} & 2+3\mathbf{i} \\ 0 & 1+\mathbf{i} & 3+2\mathbf{i} \end{vmatrix} = \begin{vmatrix} \mathbf{i} & 7+\mathbf{i} & 3-2\mathbf{i} \\ 0 & 1+\mathbf{i} & 2+3\mathbf{i} \\ 0 & 0 & 1-\mathbf{i} \end{vmatrix} = \mathbf{i}(1+\mathbf{i})(1-\mathbf{i}) = 2\mathbf{i},$$

d.h. $\det(A) = 2\mathbf{i}$. Insbesondere sehen wir, dass A invertierbar ist.

V.1.13. BEISPIEL. Liegt eine Matrix vor, die eine Zeile oder Spalte mit vielen Nullern hat, so kann es für die Bestimmung der Determinante hilfreich sein nach dieser Spalte bzw. Zeile zu entwickeln. Durch Entwicklung nach der dritten Zeile erhalten wir etwa:

$$\det \begin{pmatrix} 1 & 7 & 1 & 17 & 2 & 1 \\ 1 & 9 & 3 & 19 & 5 & 3 \\ 0 & 0 & 0 & 7 & 0 & 0 \\ -1 & -5 & 4 & 23 & 7 & 4 \\ -2 & -14 & 1 & 29 & 6 & 5 \\ 1 & 5 & -4 & 31 & -3 & 5 \end{pmatrix} = (-1)^{3+4} \cdot 7 \cdot \det \begin{pmatrix} 1 & 7 & 1 & 2 & 1 \\ 1 & 9 & 3 & 5 & 3 \\ -1 & -5 & 4 & 7 & 4 \\ -2 & -14 & 1 & 6 & 5 \\ 1 & 5 & -4 & -3 & 5 \end{pmatrix} = -7 \cdot 120 = 840.$$

wobei wir im vorletzten Gleichheitszeichen die Berechnung aus Beispiel V.1.11 verwendet haben.

V.1.14. SATZ (Vandermonde-Determinante). Für $x_0, x_1, \dots, x_n \in \mathbb{K}$ gilt

$$\det \begin{pmatrix} 1 & x_0 & x_0^2 & x_0^3 & \cdots & x_0^n \\ 1 & x_1 & x_1^2 & x_1^3 & \cdots & x_1^n \\ 1 & x_2 & x_2^2 & x_2^3 & \cdots & x_2^n \\ \vdots & \vdots & \vdots & \vdots & & \vdots \\ 1 & x_n & x_n^2 & x_n^3 & \cdots & x_n^n \end{pmatrix} = \prod_{0 \leq i < j \leq n} (x_j - x_i)$$

BEWEIS. Wir führen den Beweis durch Induktion nach n . Die Fälle $n = 0$ und $n = 1$ sind trivial. Nun zum Induktionsschritt von $n - 1$ nach n : Subtrahieren wir die vorletzte Spalte x_0 Mal von der letzten Spalte, und dann x_0 Mal die drittletzte Spalte von der vorletzten, etc. erhalten wir:

$$\begin{aligned} \det \begin{pmatrix} 1 & x_0 & x_0^2 & x_0^3 & \cdots & x_0^n \\ 1 & x_1 & x_1^2 & x_1^3 & \cdots & x_1^n \\ 1 & x_2 & x_2^2 & x_2^3 & \cdots & x_2^n \\ \vdots & \vdots & \vdots & \vdots & & \vdots \\ 1 & x_n & x_n^2 & x_n^3 & \cdots & x_n^n \end{pmatrix} &= \det \begin{pmatrix} 1 & 0 & 0 & 0 & \cdots & 0 \\ 1 & x_1 - x_0 & (x_1 - x_0)x_1 & (x_1 - x_0)x_1^2 & \cdots & (x_1 - x_0)x_1^{n-1} \\ 1 & x_2 - x_0 & (x_2 - x_0)x_2 & (x_2 - x_0)x_2^2 & \cdots & (x_2 - x_0)x_2^{n-1} \\ \vdots & \vdots & \vdots & \vdots & & \vdots \\ 1 & x_n - x_0 & (x_n - x_0)x_n & (x_n - x_0)x_n^2 & \cdots & (x_n - x_0)x_n^{n-1} \end{pmatrix} \\ &= \det \begin{pmatrix} x_1 - x_0 & (x_1 - x_0)x_1 & (x_1 - x_0)x_1^2 & \cdots & (x_1 - x_0)x_1^{n-1} \\ x_2 - x_0 & (x_2 - x_0)x_2 & (x_2 - x_0)x_2^2 & \cdots & (x_2 - x_0)x_2^{n-1} \\ \vdots & \vdots & \vdots & & \vdots \\ x_n - x_0 & (x_n - x_0)x_n & (x_n - x_0)x_n^2 & \cdots & (x_n - x_0)x_n^{n-1} \end{pmatrix} \\ &= \prod_{0 \leq i < j \leq n} (x_j - x_i) \det \begin{pmatrix} 1 & x_1 & x_1^2 & \cdots & x_1^{n-1} \\ 1 & x_2 & x_2^2 & \cdots & x_2^{n-1} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & x_n & x_n^2 & \cdots & x_n^{n-1} \end{pmatrix} \\ &= \prod_{0 \leq i < j \leq n} (x_j - x_i) \prod_{1 \leq i < j \leq n} (x_j - x_i) = \prod_{0 \leq i < j \leq n} (x_j - x_i) \end{aligned}$$

wobei wir im vorletzten Gleichheitszeichen die Induktionsvoraussetzung verwendet haben. \square

V.1.15. BEMERKUNG. Aus dem vorangehenden Satz erhalten wir erneut, dass die Vandermonde-Matrix für paarweise verschiedene x_0, x_1, \dots, x_n invertierbar ist, vgl. Satz IV.6.26.

V.2. Determinanten und Gleichungssysteme. Mit Hilfe von Determinanten lassen sich die Lösungen eines eindeutig lösbaren Gleichungssystems explizit angeben:

V.2.1. SATZ (Cramer'sche Regel). Sei \mathbb{K} ein Körper, $n \in \mathbb{N}$, $A \in M_{n \times n}(\mathbb{K})$ eine invertierbare Matrix und $y \in \mathbb{K}^n$. Dann besitzt das Gleichungssystem

$$Ax = y$$

eine eindeutige Lösung $x \in \mathbb{K}^n$ und für die i -te Komponente von x gilt

$$x_i = \frac{1}{\det(A)} \det(a_1 | \cdots | a_{i-1} | y | a_{i+1} | \cdots | a_n),$$

wobei $A = (a_1 | \cdots | a_n)$ die Spalten von A bezeichnen, $1 \leq i \leq n$.

BEWEIS. Da A invertierbar ist, existiert ein eindeutiges $x \in \mathbb{K}^n$ mit $Ax = y$. Sei nun $i \in \{1, \dots, n\}$ fix und betrachte die $(n \times n)$ -Matrix

$$X = (e_1 | \cdots | e_{i-1} | x | e_{i+1} | \cdots | e_n).$$

Dann gilt

$$AX = (Ae_1 | \cdots | Ae_{i-1} | Ax | Ae_{i+1} | \cdots | Ae_n) = (a_1 | \cdots | a_{i-1} | y | a_{i+1} | \cdots | a_n)$$

und nach Satz V.1.3(i) daher

$$\det(A) \det(X) = \det(AX) = \det(a_1 | \cdots | a_{i-1} | y | a_{i+1} | \cdots | a_n).$$

Mit den Rechenregeln in Satz V.1.3 folgt $\det(X) = x_i$ und damit der Satz. \square

V.2.2. BEISPIEL. Wir wollen das Gleichungssystem

$$\begin{pmatrix} 1 & 2 & 5 \\ -2 & 3 & -1 \\ 3 & 1 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 8 \\ 5 \\ 9 \end{pmatrix}$$

mit Hilfe der Cramer'schen Regel lösen. Dazu berechnen wir folgende Determinanten, vgl. Beispiel V.1.6:

$$\det(A) = \begin{vmatrix} 1 & 2 & 5 \\ -2 & 3 & -1 \\ 3 & 1 & 1 \end{vmatrix} = 3 - 6 - 10 - 45 + 1 + 4 = -53$$

$$\det(y|a_2|a_3) = \begin{vmatrix} 8 & 2 & 5 \\ 5 & 3 & -1 \\ 9 & 1 & 1 \end{vmatrix} = 24 - 18 + 25 - 135 + 8 - 10 = -106$$

$$\det(a_1|y|a_3) = \begin{vmatrix} 1 & 8 & 5 \\ -2 & 5 & -1 \\ 3 & 9 & 1 \end{vmatrix} = 5 - 24 - 90 - 75 + 9 + 16 = -159$$

$$\det(a_1|a_2|y) = \begin{vmatrix} 1 & 2 & 8 \\ -2 & 3 & 5 \\ 3 & 1 & 9 \end{vmatrix} = 27 + 30 - 16 - 72 - 5 + 36 = 0$$

Nach Satz V.2.1 gilt daher

$$x = \frac{1}{-53} \begin{pmatrix} -106 \\ -159 \\ 0 \end{pmatrix} = \begin{pmatrix} 2 \\ 3 \\ 0 \end{pmatrix}. \quad (\text{V.3})$$

Zum Vergleich wollen wir das selbe Gleichungssystem nochmals mit Hilfe des in Abschnitt IV.4 beschriebenen Algorithmus lösen:

$$\begin{aligned} \left(\begin{array}{ccc|c} 1 & 2 & 5 & 8 \\ -2 & 3 & -1 & 5 \\ 3 & 1 & 1 & 9 \end{array} \right) &\rightsquigarrow \left(\begin{array}{ccc|c} 1 & 2 & 5 & 8 \\ 0 & 7 & 9 & 21 \\ 0 & -5 & -14 & -15 \end{array} \right) &\rightsquigarrow \left(\begin{array}{ccc|c} 1 & 2 & 5 & 8 \\ 0 & 1 & 9/7 & 3 \\ 0 & -5 & -14 & -15 \end{array} \right) \\ &\rightsquigarrow \left(\begin{array}{ccc|c} 1 & 2 & 5 & 8 \\ 0 & 1 & 9/7 & 3 \\ 0 & 0 & -63/7 & 0 \end{array} \right) &\rightsquigarrow \left(\begin{array}{ccc|c} 1 & 2 & 5 & 8 \\ 0 & 1 & 9/7 & 3 \\ 0 & 0 & 1 & 0 \end{array} \right) \\ &&&&\rightsquigarrow \left(\begin{array}{ccc|c} 1 & 2 & 0 & 8 \\ 0 & 1 & 0 & 3 \\ 0 & 0 & 1 & 0 \end{array} \right) &\rightsquigarrow \left(\begin{array}{ccc|c} 1 & 0 & 0 & 2 \\ 0 & 1 & 0 & 3 \\ 0 & 0 & 1 & 0 \end{array} \right) \end{aligned}$$

Wir erhalten so die selbe Lösung (V.3).

Mit Hilfe von Determinanten können wir auch eine explizite Formel für die Inverse einer Matrix angeben:

V.2.3. SATZ. Sei \mathbb{K} ein Körper, $n \in \mathbb{N}$, $A \in M_{n \times n}(\mathbb{K})$ und bezeichne \tilde{A} die Matrix mit Einträgen $\tilde{A}_{ji} := (-1)^{i+j} \det(\hat{A}_{(ij)})$, d.h.

$$\tilde{A} = \begin{pmatrix} +\det(\hat{A}_{(11)}) & -\det(\hat{A}_{(12)}) & +\det(\hat{A}_{(13)}) & \cdots \\ -\det(\hat{A}_{(21)}) & +\det(\hat{A}_{(22)}) & -\det(\hat{A}_{(23)}) & \cdots \\ +\det(\hat{A}_{(31)}) & -\det(\hat{A}_{(32)}) & +\det(\hat{A}_{(33)}) & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}^t$$

wobei $\hat{A}_{(ij)}$ jene $(n-1) \times (n-1)$ -Matrix bezeichnet, die wir aus A durch Streichen der i -ten Zeile und j -ten Spalte erhalten. Dann gilt

$$A\tilde{A} = \det(A)I_n.$$

Ist A invertierbar, so erhalten wir folgende explizite Formel für die Inverse:

$$A^{-1} = \frac{1}{\det(A)}\tilde{A}.$$

BEWEIS. Entwickeln nach der i -ten Zeile, siehe Satz V.1.3(1), liefert

$$\det(A) = \sum_{j=1}^n (-1)^{i+j} A_{ij} \det(\hat{A}_{(ij)}) = \sum_{j=1}^n A_{ij} \tilde{A}_{ji} = (A\tilde{A})_{ii}.$$

Die Matrizen $A\tilde{A}$ und $\det(A)I_n$ stimmen daher längs der Diagonale überein. Sei nun $k \neq i$ und bezeichne A' die Matrix die wir erhalten, wenn wir die i -te Zeile

von A durch die k -te Zeile von A ersetzen. Dann hat A' zwei gleiche Zeilen, es gilt daher $\det(A') = 0$. Nach Konstruktion gilt weiters $A'_{ij} = A_{kj}$. Schließlich haben wir $\hat{A}'_{(ij)} = \hat{A}_{(ij)}$, denn A' und A unterscheiden sich nur in der i -ten Zeile, und diese wird ja gestrichen. Entwickeln nach der i -ten Zeile liefert somit:

$$\begin{aligned} 0 = \det(A') &= \sum_{j=1}^n (-1)^{i+j} A'_{ij} \det(\hat{A}'_{(ij)}) \\ &= \sum_{j=1}^n (-1)^{i+j} A_{kj} \det(\hat{A}_{(ij)}) = \sum_{j=1}^n A_{kj} \tilde{A}_{ji} = (A\tilde{A})_{ki}. \end{aligned}$$

Die Matrizen $A\tilde{A}$ und $\det(A)I_n$ stimmen daher auch außerhalb der Diagonale überein. Dies zeigt $A\tilde{A} = \det(A)I_n$. Im invertierbaren Fall ist die Inverse also durch $A^{-1} = \det(A)^{-1}\tilde{A}$ gegeben. \square

V.2.4. BEMERKUNG. Für die Inverse einer invertierbaren (2×2) -Matrix erhalten wir aus Satz V.2.3 erneut, siehe Beispiel II.4.8, die Formel

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}^{-1} = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

V.2.5. BEMERKUNG. Für die Inverse einer invertierbaren (3×3) -Matrix erhalten wir aus Satz V.2.3 die Formel:

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}^{-1} = \frac{1}{\det(A)} \begin{pmatrix} + | \begin{smallmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{smallmatrix} | & - | \begin{smallmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{smallmatrix} | & + | \begin{smallmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{smallmatrix} | \\ - | \begin{smallmatrix} a_{12} & a_{13} \\ a_{32} & a_{33} \end{smallmatrix} | & + | \begin{smallmatrix} a_{11} & a_{13} \\ a_{31} & a_{33} \end{smallmatrix} | & - | \begin{smallmatrix} a_{11} & a_{12} \\ a_{31} & a_{32} \end{smallmatrix} | \\ + | \begin{smallmatrix} a_{12} & a_{13} \\ a_{22} & a_{23} \end{smallmatrix} | & - | \begin{smallmatrix} a_{11} & a_{13} \\ a_{21} & a_{23} \end{smallmatrix} | & + | \begin{smallmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{smallmatrix} | \end{pmatrix}^t$$

Damit erhalten wir etwa:

$$\begin{aligned} \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 2 \\ 0 & 2 & 1 \end{pmatrix}^{-1} &= \frac{1}{7} \begin{pmatrix} + | \begin{smallmatrix} 3 & 2 \\ 2 & 1 \end{smallmatrix} | & - | \begin{smallmatrix} 2 & 2 \\ 0 & 1 \end{smallmatrix} | & + | \begin{smallmatrix} 2 & 3 \\ 0 & 2 \end{smallmatrix} | \\ - | \begin{smallmatrix} 2 & 3 \\ 2 & 1 \end{smallmatrix} | & + | \begin{smallmatrix} 1 & 3 \\ 0 & 1 \end{smallmatrix} | & - | \begin{smallmatrix} 1 & 2 \\ 0 & 2 \end{smallmatrix} | \\ + | \begin{smallmatrix} 2 & 3 \\ 3 & 2 \end{smallmatrix} | & - | \begin{smallmatrix} 1 & 3 \\ 2 & 2 \end{smallmatrix} | & + | \begin{smallmatrix} 1 & 2 \\ 2 & 3 \end{smallmatrix} | \end{pmatrix}^t \\ &= \frac{1}{7} \begin{pmatrix} -1 & -2 & 4 \\ 4 & 1 & -2 \\ -5 & 4 & -1 \end{pmatrix}^t = \frac{1}{7} \begin{pmatrix} -1 & 4 & -5 \\ -2 & 1 & 4 \\ 4 & -2 & -1 \end{pmatrix} \end{aligned}$$

V.3. Permutationen und Leibniz'sche Formel. Es bezeichne \mathfrak{S}_n die Menge aller *Permutationen* der Menge $\{1, \dots, n\}$, d.h. die Menge aller bijektiven Abbildungen $\sigma: \{1, \dots, n\} \rightarrow \{1, \dots, n\}$. Bezüglich der Komposition von Abbildungen bildet \mathfrak{S}_n eine Gruppe, die aus $n!$ Elementen besteht. Unter einer *Transposition* verstehen wir eine Permutation die zwei Elemente vertauscht und alle anderen fix lässt. Es bezeichne $\tau_{ij} \in \mathfrak{S}_n$ jene Transposition die die beiden Zahlen i und j vertauscht und alle anderen fix lässt, $i, j \in \{1, \dots, n\}$, $i \neq j$. Beachte, dass für jede Transposition $\tau \in \mathfrak{S}_n$ die Relation $\tau \circ \tau = 1$ gilt.

Wir beginnen mit folgendem einfachen Hilfsatz, den wir eigentlich schon im Beweis von Lemma V.1.1(h) verwendet haben:

V.3.1. LEMMA. *Jede Permutation $\sigma \in \mathfrak{S}_n$ lässt sich als Produkt von Transpositionen schreiben. Dabei genügt es Transpositionen zu verwenden, die jeweils nur benachbarte Zahlen vertauschen.*

BEWEIS. Wir führen den Beweis durch Induktion nach n . Die Fälle $n = 1$ und $n = 2$ sind trivial. Für den Induktionsschritt sei nun $\sigma \in \mathfrak{S}_n$ und $i := \sigma(n)$. Nach Konstruktion lässt die Permutation

$$\tilde{\sigma} := \tau_{n-1,n} \circ \tau_{n-2,n-1} \circ \cdots \circ \tau_{i+1,i+2} \circ \tau_{i,i+1} \circ \sigma \quad (\text{V.4})$$

die Zahl n fix, d.h. $\tilde{\sigma}(n) = n$. Sie kann daher als Permutation von $\{1, \dots, n-1\}$ aufgefasst werden. Nach Induktionsvoraussetzung lässt sich $\tilde{\sigma}$ in der Form $\tilde{\sigma} = \tilde{\tau}_1 \circ \cdots \circ \tilde{\tau}_N$ schreiben, wobei jedes $\tilde{\tau}_j$ eine Transposition bezeichnet, die zwei benachbarte Zahlen vertauscht. Mit (V.4) erhalten wir die gesuchte Darstellung,

$$\sigma = \tau_{i,i+1} \circ \cdots \circ \tau_{n-1,n} \circ \tilde{\sigma} = \tau_{i,i+1} \circ \cdots \circ \tau_{n-1,n} \circ \tilde{\tau}_1 \circ \cdots \circ \tilde{\tau}_N.$$

Damit ist der Induktionsschritt gezeigt und das Lemma bewiesen. \square

V.3.2. LEMMA. *Die Abbildung*

$$\text{sgn}: \mathfrak{S}_n \rightarrow \{1, -1\}, \quad \text{sgn}(\sigma) = (-1)^{\#\{(i,j) \mid i < j \text{ und } \sigma(i) > \sigma(j)\}},$$

ist ein Gruppenhomomorphismus, der jede Transposition auf -1 abbildet. Dabei fassen wir $\{1, -1\}$ bezüglich Multiplikation als (abelsche) Gruppe auf. Jeder weitere Gruppenhomomorphismus $\mathfrak{S}_n \rightarrow \{1, -1\}$ mit dieser Eigenschaft muss mit sgn übereinstimmen.

BEWEIS. Sei $\sigma \in \mathfrak{S}_n$ und $\tau = \tau_{k,k+1}$, wobei $1 \leq k < n$. Dann gilt

$$\#\{(i, j) \mid i < j, (\sigma \circ \tau)(i) > (\sigma \circ \tau)(j)\} = \#\{(i, j) \mid i < j, \sigma(i) > \sigma(j)\} \pm 1$$

je nachdem, ob $\sigma(k) > \sigma(k+1)$ oder $\sigma(k) < \sigma(k+1)$. Jedenfalls folgt

$$\text{sgn}(\sigma \circ \tau) = -\text{sgn}(\sigma). \quad (\text{V.5})$$

Sei nun $\sigma' \in \mathfrak{S}_n$. Nach Lemma V.3.1 lässt sich σ' in der Form $\sigma' = \tau'_N \circ \cdots \circ \tau'_1$ schreiben, wobei jedes τ'_i eine Transposition bezeichnet, die zwei benachbarte Zahlen vertauscht. Induktiv erhalten wir aus (V.5)

$$\begin{aligned} \text{sgn}(\sigma \circ \sigma') &= \text{sgn}(\sigma \circ \tau'_N \circ \cdots \circ \tau'_1) \\ &= -\text{sgn}(\sigma \circ \tau'_N \circ \cdots \circ \tau'_2) = \cdots = (-1)^N \text{sgn}(\sigma) \end{aligned}$$

und analog

$$\begin{aligned} \text{sgn}(\sigma') &= \text{sgn}(\tau'_N \circ \cdots \circ \tau'_1) \\ &= -\text{sgn}(\tau'_N \circ \cdots \circ \tau'_2) = \cdots = (-1)^N \text{sgn}(\text{id}) = (-1)^N, \end{aligned}$$

denn $\text{sgn}(\text{id}) = 1$, da ja $\#\{(i, j) \mid i < j, \text{id}(i) > \text{id}(j)\} = 0$. Dies zeigt

$$\text{sgn}(\sigma \circ \sigma') = \text{sgn}(\sigma) \text{sgn}(\sigma'),$$

für beliebige $\sigma, \sigma' \in \mathfrak{S}_n$, also ist $\text{sgn}: \mathfrak{S}_n \rightarrow \{1, -1\}$ ein Homomorphismus. Sei nun $\tau = \tau_{k,l}$ eine beliebige Transposition, $k \neq l$. Dann existiert $\sigma \in \mathfrak{S}_n$, sodass $\sigma(1) = k$ und $\sigma(2) = l$. Es folgt $\tau = \sigma \circ \tau_{12} \circ \sigma^{-1}$, und daher

$$\text{sgn}(\tau) = \text{sgn}(\sigma) \text{sgn}(\tau_{12}) \text{sgn}(\sigma^{-1}) = \text{sgn}(\sigma) \text{sgn}(\tau_{12}) \text{sgn}(\sigma)^{-1} = \text{sgn}(\tau_{12}) = -1.$$

Die Eindeutigkeitsaussage folgt sofort aus Lemma V.3.1. \square

V.3.3. BEMERKUNG. Der Kern des Homomorphismus $\text{sgn}: \mathfrak{S}_n \rightarrow \{1, -1\}$,

$$A_n := \{\sigma \in \mathfrak{S}_n : \text{sgn}(\sigma) = 1\},$$

bildet eine Untergruppe von \mathfrak{S}_n , sie wird die *alternierende Gruppe* genannt.

V.3.4. BEMERKUNG. Nach Lemma V.3.1 lässt sich jede Permutation $\sigma \in \mathfrak{S}_n$ als Produkt von Transpositionen schreiben, $\sigma = \tau_N \circ \cdots \circ \tau_1$. Diese Darstellung ist allerdings nicht eindeutig, etwa haben wir auch $\sigma = \tau_{12} \circ \tau_{12} \circ \tau_N \circ \cdots \circ \tau_1$. Nach Lemma V.3.2 gilt jedoch $\text{sgn}(\sigma) = (-1)^N$, d.h. die Parität der Anzahl der Faktoren ist unabhängig von der Darstellung. Für $\text{sgn}(\sigma) = 1$ muss N gerade sein, d.h. *jede* Darstellung von σ als Produkt von Transpositionen hat eine gerade Anzahl von Faktoren. Solche Permutationen werden als *gerade Permutationen* bezeichnet. Ist $\text{sgn}(\sigma) = -1$, dann muss N ungerade sein, d.h. *jede* Darstellung von σ als Produkt von Transpositionen hat eine ungerade Anzahl von Faktoren. Solche Permutationen werden *ungerade Permutationen* genannt.

V.3.5. BEISPIEL. Die Permutation $\sigma \in \mathfrak{S}_3$, $\sigma(1) = 2$, $\sigma(2) = 3$, $\sigma(3) = 1$, lässt sich in der Form $\sigma = \tau_{12} \circ \tau_{23}$ schreiben, es gilt daher $\text{sgn}(\sigma) = 1$.

V.3.6. SATZ (Leibniz'sche Formel). Sei \mathbb{K} ein Körper und $n \in \mathbb{N}$. Dann gilt für jede Matrix $A \in M_{n \times n}(\mathbb{K})$

$$\det(A) = \sum_{\sigma \in \mathfrak{S}_n} \text{sgn}(\sigma) A_{1,\sigma(1)} A_{2,\sigma(2)} \cdots A_{n,\sigma(n)}.$$

BEWEIS. Wir bezeichnen die rechte Seite der Formel mit

$$\delta: M_{n \times n}(\mathbb{K}) \rightarrow \mathbb{K}, \quad \delta(A) := \sum_{\sigma \in \mathfrak{S}_n} \text{sgn}(\sigma) A_{1,\sigma(1)} A_{2,\sigma(2)} \cdots A_{n,\sigma(n)}. \quad (\text{V.6})$$

Offensichtlich ist $\delta(A)$ linear in jeder Spalte von A . Dies gilt sogar für jeden Ausdruck der Form $A_{1,\sigma(1)} A_{2,\sigma(2)} \cdots A_{n,\sigma(n)}$, denn er enthält aus jeder Spalte von A genau einen Faktor. Stimmen die Spalten k und l von A überein, $k \neq l$, und bezeichnet $\tau = \tau_{kl}$ die Transposition, die k mit l vertauscht, dann gilt $A_{i,j} = A_{i,\tau(j)}$ und daher auch $A_{i,\sigma(i)} = A_{i,(\tau \circ \sigma)(i)}$, für alle $i, j \in \{1, \dots, n\}$. Die beiden Summanden in (V.6), die den Permutationen σ und $\tau \circ \sigma$ entsprechen, unterscheiden sich also nur durch ihr Vorzeichen, $\text{sgn}(\tau \circ \sigma) = -\text{sgn}(\sigma)$, und kürzen einander. Dies zeigt, dass $\delta(A) = 0$ gilt, wenn zwei Spalten von A übereinstimmen. Schließlich

gilt auch $\delta(I_n) = 1$, denn für $A = I_n$ liefert nur $\sigma = \text{id}$ einen nicht-trivialen Beitrag in der Summe (V.6), nämlich 1. Aus der Eindeutigkeitsaussage in Satz V.1.3 folgt daher $\det(A) = \delta(A)$. \square

V.3.7. BEMERKUNG. Der Beweis von Satz V.3.6 liefert eine neue Konstruktion der Determinantenfunktion, die unabhängig von der im Beweis des Satzes V.1.3 besprochenen ist.

V.3.8. BEISPIEL. Für (3×3) -Matrizen erhalten wir:

σ	$\sigma(1)$	$\sigma(2)$	$\sigma(3)$	$\text{sgn}(\sigma)$	Term in Leibniz' Formel
id	1	2	3	1	$a_{11}a_{22}a_{33}$
τ_{12}	2	1	3	-1	$-a_{12}a_{21}a_{33}$
τ_{13}	3	2	1	-1	$-a_{13}a_{22}a_{31}$
τ_{23}	1	3	2	-1	$-a_{11}a_{23}a_{32}$
$\tau_{12} \circ \tau_{23}$	2	3	1	1	$a_{12}a_{23}a_{31}$
$\tau_{23} \circ \tau_{12}$	3	1	2	1	$a_{13}a_{21}a_{32}$

Nach Satz V.3.6 gilt daher

$$\det \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} = a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33} - a_{13}a_{22}a_{31}.$$

Wir erhalten also erneut die Regel von Sarrus, vgl. Bemerkung V.1.6.

V.4. Determinanten von Endomorphismen. Sei V ein endlich-dimensionaler \mathbb{K} -Vektorraum und $\varphi: V \rightarrow V$ linear. Sind B und \tilde{B} zwei geordnete Basen von V , dann gilt $[\varphi]_{\tilde{B}\tilde{B}} = T[\varphi]_{BB}T^{-1}$, wobei $T = T_{\tilde{B}B} \in \text{GL}_n(\mathbb{K})$ und $n = \dim(V)$, siehe Korollar IV.6.21. Mit Satz V.1.3 folgt

$$\det([\varphi]_{\tilde{B}\tilde{B}}) = \det(T[\varphi]_{\tilde{B}\tilde{B}}T^{-1}) = \det(T) \det([\varphi]_{BB}) \det(T)^{-1} = \det([\varphi]_{BB}),$$

d.h. die Determinanten der Matrixdarstellung von φ ist unabhängig von der Basis. Dies erlaubt es eine Determinante für Endomorphismen $\varphi: V \rightarrow V$ zu definieren:

V.4.1. DEFINITION (Determinante eines Endomorphismus). Sei V ein endlich-dimensionaler \mathbb{K} -Vektorraum und $\varphi: V \rightarrow V$ linear. Unter der Determinante von φ verstehen wir die Zahl

$$\det(\varphi) := \det([\varphi]_{BB}),$$

wobei B eine geordnete Basis von V bezeichnet. Nach den Ausführungen oben hängt dies nicht von der Wahl der Basis B ab, und ist daher wohldefiniert.

V.4.2. KOROLLAR. Sei V ein endlich-dimensionaler \mathbb{K} -Vektorraum und

$$\det: \text{end}(V) \rightarrow \mathbb{K}.$$

die eben definierte Abbildung, siehe Definition V.4.1. Dann gilt

$$\det(\psi \circ \varphi) = \det(\psi) \det(\varphi) \quad \text{und} \quad \det(\text{id}_V) = 1,$$

für alle $\varphi, \psi \in \text{end}(V)$. Weiters ist $\varphi \in \text{end}(V)$ genau dann invertierbar, wenn $\det(\varphi) \neq 0$, und in diesem Fall gilt

$$\det(\varphi^{-1}) = \det(\varphi)^{-1}.$$

Die Determinante schränkt sich daher zu einem surjektiven Gruppenhomomorphismus

$$\det: \text{GL}(V) \rightarrow \mathbb{K} \setminus \{0\}$$

ein, deren Kern, $\text{SL}(V) := \{\varphi \in \text{end}(V) : \det(\varphi) = 1\}$, eine Untergruppe von $\text{GL}(V)$ bildet. Schließlich haben wir

$$\det(\varphi^t) = \det(\varphi),$$

für jedes $\varphi \in \text{end}(V)$, wobei $\varphi^t \in \text{end}(V^*)$ die zu φ duale Abbildung bezeichnet.

BEWEIS. Es sei B eine geordnete Basis von V . Nach Korollar IV.6.21 gilt $[\psi \circ \varphi]_{BB} = [\psi]_{BB}[\varphi]_{BB}$ und mit Satz V.1.3(i) daher

$$\begin{aligned} \det(\psi \circ \varphi) &= \det([\psi \circ \varphi]_{BB}) = \det([\psi]_{BB}[\varphi]_{BB}) \\ &= \det([\psi]_{BB}) \det([\varphi]_{BB}) = \det(\psi) \det(\varphi). \end{aligned}$$

Da $[\text{id}_V]_{BB} = I_n$, erhalten wir analog $\det(\text{id}_V) = \det([\text{id}_V]_{BB}) = \det(I_n) = 1$. Nach Korollar IV.6.21 ist φ genau dann invertierbar, wenn $[\varphi]_{BB}$ eine invertierbare Matrix ist. Nach Satz V.1.3(j) ist dies genau dann der Fall, wenn $\det(\varphi) = \det([\varphi]_{BB}) \neq 0$ gilt. Im invertierbaren Fall folgt

$$1 = \det(\text{id}_V) = \det(\varphi^{-1} \circ \varphi) = \det(\varphi^{-1}) \det(\varphi),$$

also $\det(\varphi^{-1}) = \det(\varphi)^{-1}$. Nach Korollar IV.6.21 gilt $[\varphi^t]_{B^*B^*} = [\varphi]_{BB}^t$, also

$$\det(\varphi^t) = \det([\varphi^t]_{B^*B^*}) = \det([\varphi]_{BB}^t) = \det([\varphi]_{BB}) = \det(\varphi),$$

wobei wir beim vorletzten Gleichheitszeichen Satz V.1.3(k) verwendet haben. \square

V.4.3. BEISPIEL. Bezeichne $V = \mathbb{R}[z]_{\leq 3}$ den Vektorraum der reellen Polynome vom Grad kleiner gleich 3 und betrachte die lineare Abbildung $\varphi: V \rightarrow V$, $\varphi(p) := p' - p$. Wir wollen die Determinante von φ bestimmen. Dazu wählen wir eine geordnete Basis B von V , etwa $p_0 = 1$, $p_1 = z$, $p_2 = z^2$, $p_3 = z^3$. Dann folgt

$$[\varphi]_{BB} = \begin{pmatrix} -1 & 1 & 0 & 0 \\ 0 & -1 & 2 & 0 \\ 0 & 0 & -1 & 3 \\ 0 & 0 & 0 & -1 \end{pmatrix}$$

also $\det(\varphi) = \det([\varphi]_{BB}) = (-1)^4 = 1$.

V.5. Orientierung reeller Vektorräume. Die Euklidische Ebene \mathbb{R}^2 besitzt zwei Orientierungen: im Uhrzeigersinn oder gegen den Uhrzeigersinn. Auch auf \mathbb{R}^3 gibt es zwei Orientierungen: Korkenzieher für Links- bzw. Rechtshänder. Wir wollen diesen eher vagen Begriff der Orientierung nun präzisieren und auf beliebige Dimensionen ausweiten. Für den Rest dieses Abschnitts sei V ein endlichdimensionaler reeller Vektorraum und $n := \dim(V) \geq 1$.

Wir definieren auf der Menge der geordneten Basen von V eine Relation:

$$B \sim B' :\Leftrightarrow \det(T_{B'B}) > 0.$$

Dabei sind B und B' zwei geordnete Basen von V und $T_{B'B}$ bezeichnet die Matrix zum Basiswechsel von B nach B' , siehe Abschnitt IV.6.

V.5.1. LEMMA. *Dies ist eine Äquivalenzrelation.*

BEWEIS. Die Relation ist reflexiv, denn wegen $\det(T_{BB}) = \det(I_n) = 1 > 0$ gilt $B \sim B$. Die Relation ist symmetrisch, denn aus $B \sim B'$ folgt $\det(T_{B'B}) > 0$, also auch $\det(T_{BB'}) = \det(T_{B'B}^{-1}) = \det(T_{B'B})^{-1} > 0$ und damit $B' \sim B$. Gilt $B \sim B'$ und $B' \sim B''$, d.h. $\det(T_{B'B}) > 0$ und $\det(T_{B''B'}) > 0$, dann folgt $\det(T_{B''B}) = \det(T_{B''B'}T_{B'B}) = \det(T_{B''B'})\det(T_{B'B}) > 0$, also $B \sim B''$. Dies zeigt, dass die Relation auch transitiv ist. \square

Wir bezeichnen die Menge der Äquivalenzklassen dieser Relation mit $\mathcal{O}(V)$. Die Elemente von $\mathcal{O}(V)$ werden *Orientierungen* von V genannt. Ist B eine geordnete Basis von V , dann bezeichne $\mathfrak{o}_B \in \mathcal{O}(V)$, die von B repräsentierte Orientierung (Äquivalenzklasse). Zwei geordnete Basen B und B' werden *gleich-orientiert* genannt, wenn sie dieselbe Orientierung repräsentieren, d.h. wenn $\mathfrak{o}_B = \mathfrak{o}_{B'}$ gilt. Nach Definition ist dies genau dann der Fall, wenn $\det(T_{B'B}) > 0$ gilt.

V.5.2. LEMMA. *Die Menge $\mathcal{O}(V)$ besteht aus genau zwei Elementen, d.h. V besitzt genau zwei Orientierungen. Ist $B = (b_1, \dots, b_n)$ eine geordnete Basis von V und $\mathfrak{o}_B \in \mathcal{O}(V)$ die von ihr repräsentierte Orientierung, dann repräsentiert die Basis $\tilde{B} = (-b_1, b_2, \dots, b_n)$ die andere Orientierung, d.h. $\mathfrak{o}_B \neq \mathfrak{o}_{\tilde{B}}$.*

BEWEIS. Die beiden Basen B und \tilde{B} sind nicht gleich-orientiert, denn

$$\det(T_{\tilde{B}B}) = \det \begin{pmatrix} -1 & & \\ & I_{n-1} & \\ & & \end{pmatrix} = -1 < 0.$$

In anderen Worten $\mathfrak{o}_B \neq \mathfrak{o}_{\tilde{B}}$. Die Menge $\mathcal{O}(V)$ besitzt daher wenigstens zwei verschiedene Elemente. Sei nun B' eine weitere geordnete Basis von V und $\mathfrak{o}_{B'} \neq \mathfrak{o}_B$, d.h. $\det(T_{B'B}) \leq 0$. Da $T_{B'B}$ invertierbar ist, folgt $\det(T_{B'B}) < 0$ und daher $\det(T_{B'\tilde{B}}) = \det(T_{B'B}T_{B\tilde{B}}) = \det(T_{B'B})\det(T_{B\tilde{B}}) > 0$, also $\mathfrak{o}_{B'} = \mathfrak{o}_{\tilde{B}}$. Dies zeigt, dass V genau zwei Orientierungen besitzt, nämlich \mathfrak{o}_B und $\mathfrak{o}_{\tilde{B}}$. \square

Ist $\mathfrak{o} \in \mathcal{O}(V)$ eine Orientierung von V , dann bezeichne $-\mathfrak{o}$ die andere Orientierung von V . Mit der Notation aus Lemma V.5.2 gilt daher $-\mathfrak{o}_B = \mathfrak{o}_{\tilde{B}}$.

Unter einem *orientierten Vektorraum* verstehen wir einen Vektorraum zusammen mit einer Orientierung, d.h. ein Paar (V, \mathfrak{o}) , wobei V ein endlich-dimensionaler reeller Vektorraum ist und $\mathfrak{o} \in \mathcal{O}(V)$. In dieser Situation werden die Basen in \mathfrak{o} als *positiv orientierte Basen* bezeichnet, die Basen in $-\mathfrak{o}$ werden negativ orientierte Basen genannt.

V.5.3. BEISPIEL. Unter der *Standardorientierung* von \mathbb{R}^n verstehen wir die von der Standardbasis $E = (e_1, \dots, e_n)$ repräsentierte Orientierung \mathfrak{o}_E . Bezüglich der Standardorientierung von \mathbb{R}^3 ist die Basis

$$b_1 = \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix}, \quad b_2 = \begin{pmatrix} 2 \\ -1 \\ 1 \end{pmatrix}, \quad b_3 = \begin{pmatrix} 3 \\ 2 \\ -7 \end{pmatrix}$$

positiv orientiert, denn

$$\det(T_{EB}) = \det \begin{pmatrix} 1 & 2 & 3 \\ 2 & -1 & 2 \\ 0 & 1 & 1 \end{pmatrix} = 31 > 0,$$

wobei $E = (e_1, e_2, e_3)$ die Standardbasis von \mathbb{R}^3 bezeichnet.

V.5.4. LEMMA. *Ist $\varphi: V \rightarrow W$ ein linearer Isomorphismus zwischen endlich-dimensionalen reellen Vektorräumen, und sind $B \sim B'$ zwei gleichorientierte Basen von V , dann sind auch $\varphi(B)$ und $\varphi(B')$ gleichorientierte Basen von W . Wir erhalten daher eine induzierte Bijektion*

$$\varphi: \mathcal{O}(V) \rightarrow \mathcal{O}(W), \quad \varphi(\mathfrak{o}_B) := \mathfrak{o}_{\varphi(B)}.$$

Dabei bezeichnet $\varphi(B)$ die Basis $\varphi(b_1), \dots, \varphi(b_n)$ von W , falls $B = (b_1, \dots, b_n)$.

BEWEIS. Seien also B und B' gleichorientiert, d.h. $\det(T_{B'B}) > 0$. Offensichtlich gilt $[\varphi]_{\varphi(B)B} = I_n = [\varphi]_{\varphi(B')B'}$. Somit ist

$$\begin{aligned} T_{\varphi(B')\varphi(B)} &= [\text{id}_W]_{\varphi(B')\varphi(B)} \\ &= [\varphi \circ \text{id}_V \circ \varphi^{-1}]_{\varphi(B')\varphi(B)} \\ &= [\varphi]_{\varphi(B')B'} [\text{id}_V]_{B'B} [\varphi^{-1}]_{B\varphi(B)} \\ &= [\varphi]_{\varphi(B')B'} [\text{id}_V]_{B'B} [\varphi]_{\varphi(B)B}^{-1} \\ &= I_n T_{B'B} I_n = T_{B'B} \end{aligned}$$

also $\det(T_{\varphi(B')\varphi(B)}) = \det(T_{B'B}) > 0$, d.h. die Basen $\varphi(B)$ und $\varphi(B')$ sind gleichorientiert. \square

Für einen linearen Isomorphismus $\varphi: V \rightarrow V$ tritt genau einer der folgenden beiden Fälle ein:

- (a) $\det(\varphi) > 0$: In diesem Fall gilt $\varphi(\mathfrak{o}) = \mathfrak{o}$ für alle $\mathfrak{o} \in \mathcal{O}(V)$, und φ wird *orientierungsbewahrend* genannt.

(b) $\det(\varphi) < 0$: In diesem Fall gilt $\varphi(\mathfrak{o}) = -\mathfrak{o}$ für alle $\mathfrak{o} \in \mathcal{O}(V)$ und φ wird *orientierungsumkehrend* genannt.

Dies folgt aus der Gleichung $[\varphi]_{BB} = T_{B\varphi(B)}$, wobei B eine Basis von V bezeichnet.

V.5.5. BEISPIEL. Die Rotationen $\rho: \mathbb{R}^2 \rightarrow \mathbb{R}^2$,

$$\rho \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$$

ist für jedes $\alpha \in \mathbb{R}$ orientierungsbewahrend, denn $\begin{vmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{vmatrix} = 1 > 0$.

V.5.6. BEISPIEL. Seien W und W' zwei komplementäre Teilräume von V , d.h. $V = W \oplus W'$. Bezeichnet $\sigma: V \rightarrow V$ die Spiegelung an W längs W' , dann gilt $\det(\sigma) = (-1)^{\dim(W')}$, siehe Aufgabe 24. Diese Spiegelung ist also orientierungserhaltend falls W' gerade Dimension hat, und sie ist orientierungsumkehrend, wenn W' ungerade Dimension hat. Insbesondere ist jede Spiegelung an einer Hyperebene orientierungsumkehrend. Auch folgt daraus, dass die Spiegelung am Ursprung, $-\text{id}_V: V \rightarrow V$, genau dann orientierungsbewahrend ist, wenn V gerade Dimension hat, anderenfalls ist sie orientierungsumkehrend.

V.5.7. BEMERKUNG. Die Menge der Orientierungsbewahrenden Isomorphismen bildet eine Untergruppe von $\text{GL}(V)$, siehe Aufgabe 26.

Sei $I \subseteq \mathbb{R}$ ein Intervall. Unter einer Kurve in V verstehen wir eine Abbildung $v: I \rightarrow V, t \mapsto v(t)$. Eine solche Kurve wird stetig genannt, falls ihre Koordinaten bezüglich einer geordneten Basis B von V stetig sind, d.h. die Abbildung $I \rightarrow \mathbb{R}^n, t \mapsto [v(t)]_B$, stetig ist. Dies bleibt dann für jede weitere Basis B' von V richtig, denn mit $[v(t)]_B$ ist auch $[v(t)]_{B'} = T_{B'B}[v(t)]_B$ stetig in t .

Eine Familie geordneter Basen, $B(t) = (b_1(t), \dots, b_n(t))$, $t \in I$, wird stetig genannt, falls jede der Kurven $b_i: I \rightarrow V$ stetig ist, $i = 1, \dots, n$. Dies ist genau dann der Fall, wenn jede Komponente der Basiswechselmatrix $T_{B'B(t)}$ stetig von t abhängt, wobei B' eine beliebige geordnete Basis von V bezeichnet.

V.5.8. SATZ. Sei $I \subseteq \mathbb{R}$ ein Intervall und $B(t)$, $t \in I$, eine stetige Kurve geordneter Basen von V . Dann sind je zwei Basen dieser Familie gleichorientiert, d.h. es gilt $\mathfrak{o}_{B(t_1)} = \mathfrak{o}_{B(t_2)}$, für alle $t_1, t_2 \in I$.

BEWEIS. Aufgrund der Leibniz'schen Formel ist die Abbildung

$$\omega: I \rightarrow \mathbb{R} \setminus \{0\}, \quad \omega(t) := \det(T_{B(t_1)B(t)}),$$

stetig. Weiters ist $\omega(t_1) = 1 > 0$. Aus dem Zwischenwertsatz folgt $\omega(t) > 0$, für alle $t \in I$. Es gilt somit $\mathfrak{o}_{B(t)} = \mathfrak{o}_{B(t_1)}$, für alle $t \in I$. \square

VI. Eigenwerte und Eigenvektoren

Wir wollen in diesem Abschnitt lineare Abbildungen $\varphi: V \rightarrow V$ genauer untersuchen, wobei V einen endlich-dimensionalen Vektorraum bezeichnet.

VI.1. Diagonalisierbarkeit. Wir beginnen mit der Definition einer Klasse von Endomorphismen, die besonders leicht zu verstehen sind:

VI.1.1. DEFINITION (Diagonalisierbarkeit). Sei V ein endlich-dimensionaler \mathbb{K} -Vektorraum. Eine lineare Abbildung $\varphi: V \rightarrow V$ wird *diagonalisierbar* genannt, falls eine geordnete Basis B von V existiert, sodass

$$[\varphi]_{BB} = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix},$$

wobei $n = \dim(V)$ und $\lambda_1, \dots, \lambda_n \in \mathbb{K}$.

Eine lineare Abbildung $\varphi: V \rightarrow V$ ist also genau dann diagonalisierbar, wenn eine geordnete Basis $B = (b_1, \dots, b_n)$ existiert, für die

$$\varphi(b_i) = \lambda_i b_i, \quad i = 1, \dots, n$$

gilt, wobei $\lambda_1, \dots, \lambda_n \in \mathbb{K}$ und $n = \dim(V)$.

VI.1.2. DEFINITION (Eigenwerte und Eigenvektoren). Sei V ein endlich-dimensionaler \mathbb{K} -Vektorraum und $\varphi: V \rightarrow V$ linear. Ein Skalar $\lambda \in \mathbb{K}$ wird *Eigenwert* von φ genannt, falls ein Vektor $v \in V$ existiert, sodass $v \neq 0$ und

$$\varphi(v) = \lambda v.$$

In diesem Fall wird v als *Eigenvektor* zum Eigenwert λ bezeichnet, und

$$E_\lambda := \{v \in V \mid \varphi(v) = \lambda v\} = \ker(\varphi - \lambda \text{id}_V)$$

wird der *Eigenraum* zum Eigenwert λ genannt. Unter der *geometrischen Vielfachheit* des Eigenwertes λ verstehen wir die Dimension des Eigenraums E_λ . Die Menge aller Eigenwerte wird *Spektrum* von φ genannt und mit $\sigma(\varphi)$ bezeichnet.

Eine lineare Abbildung $\varphi: V \rightarrow V$ ist also genau dann diagonalisierbar, wenn eine Basis von V existiert, die aus Eigenvektoren von φ besteht. Solche Basen werden manchmal auch als *Eigenbasen* bezeichnet. Zur Bestimmung der Eigenwerte haben wir:

VI.1.3. PROPOSITION. Sei V ein endlich-dimensionaler \mathbb{K} -Vektorraum und $\varphi: V \rightarrow V$ linear. Ein Skalar $\lambda \in \mathbb{K}$ ist genau dann Eigenwert von φ , wenn

$$\det(\varphi - \lambda \text{id}_V) = 0.$$

BEWEIS. Für $\lambda \in \mathbb{K}$ gilt:

$$\begin{aligned} \lambda \text{ ist Eigenwert von } \varphi &\Leftrightarrow \exists v \in V : \varphi(v) = \lambda v, v \neq 0 \\ &\Leftrightarrow \ker(\varphi - \lambda \text{id}_V) \neq \{0\} \\ &\Leftrightarrow \varphi - \lambda \text{id}_V : V \rightarrow V \text{ ist nicht invertierbar} \\ &\Leftrightarrow \det(\varphi - \lambda \text{id}_V) = 0 \end{aligned}$$

Dabei haben wir die beiden Korollare IV.2.11 und V.4.2 verwendet. \square

VI.1.4. BEMERKUNG. Eine lineare Abbildung $\varphi: V \rightarrow V$ ist genau dann invertierbar, wenn 0 nicht Eigenwert von φ ist. Beide Bedingungen an φ sind nämlich zu $\ker(\varphi) = \{0\}$ äquivalent, vgl. Korollar IV.2.11.

VI.1.5. BEMERKUNG. Sei $\psi: V \xrightarrow{\cong} W$ ein linearer Isomorphismus zwischen endlich-dimensionalen \mathbb{K} -Vektorräumen, und $\varphi: V \rightarrow V$ linear. Dann hat die lineare Abbildung $\psi \circ \varphi \circ \psi^{-1}: W \rightarrow W$ dieselben Eigenwerte wie φ , d.h.

$$\sigma(\psi \circ \varphi \circ \psi^{-1}) = \sigma(\varphi).$$

Ein Vektor $v \in V$ ist genau dann Eigenvektor von φ zum Eigenwert λ , wenn $\psi(v)$ Eigenvektor von $\psi \circ \varphi \circ \psi^{-1}$ zum Eigenwert λ ist, d.h. bildet b_1, \dots, b_n eine Basis von V , die aus Eigenvektoren von φ besteht, dann bildet $\psi(b_1), \dots, \psi(b_n)$ eine Basis von W , die aus Eigenvektoren von $\psi \circ \varphi \circ \psi^{-1}$ besteht. Insbesondere ist φ genau dann diagonalisierbar, wenn $\psi \circ \varphi \circ \psi^{-1}$ diagonalisierbar ist.

VI.1.6. BEMERKUNG. Sei $\varphi: V \rightarrow V$ diagonalisierbar und $B = (b_1, \dots, b_n)$ eine Basis aus Eigenvektoren,

$$[\varphi]_{BB} = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}.$$

Ist $\sigma \in \mathfrak{S}_n$ eine Permutation und bezeichnet $\tilde{B} := (b_{\sigma(1)}, \dots, b_{\sigma(n)})$ die mit σ umgeordnete Basis, dann gilt

$$[\varphi]_{\tilde{B}\tilde{B}} = \begin{pmatrix} \lambda_{\sigma(1)} & & \\ & \ddots & \\ & & \lambda_{\sigma(n)} \end{pmatrix}.$$

Die Diagonaleinträge können daher durch Übergang zu einer geeigneten Eigenbasis in beliebige Reihenfolge gebracht werden, vgl. auch Aufgabe 31.

VI.1.7. BEISPIEL. Seien W und W' komplementäre Teilräume eines endlich-dimensionalen Vektorraums V , d.h. $V = W \oplus W'$. Es bezeichne $\pi: V \rightarrow V$ die Projektion auf W längs W' . Wir wollen uns davon überzeugen, dass π diagonalisierbar ist. Seien dazu b_1, \dots, b_k eine Basis von W und b_{k+1}, \dots, b_n eine Basis

von W' . Dann bildet $B = (b_1, \dots, b_n)$ eine Basis von W und es gilt

$$[\pi]_{BB} = \begin{pmatrix} I_k & 0 \\ 0 & 0 \end{pmatrix}.$$

Der Projektor π ist daher diagonalisierbar. Die einzigen möglichen Eigenwerte sind 1 und 0, denn $\det(\pi - \lambda \text{id}_V) = \det \begin{pmatrix} (1-\lambda)I_k & 0 \\ 0 & -\lambda I_{n-k} \end{pmatrix} = (1-\lambda)^k (-\lambda)^{n-k}$, siehe Proposition VI.1.3. Gilt $W \neq \{0\} \neq W'$ so treten beide Eigenwerte tatsächlich auf, und die zugehörigen Eigenräume sind $E_1 = W$ sowie $E_0 = W'$.

Analog können wir die Spiegelung $\sigma: V \rightarrow V$ an W längs W' analysieren. In diesem Fall gilt

$$[\sigma]_{BB} = \begin{pmatrix} I_k & 0 \\ 0 & -I_{n-k} \end{pmatrix},$$

d.h. auch σ ist diagonalisierbar. Die einzigen möglichen Eigenwerte sind 1 und -1 . Gilt $W \neq \{0\} \neq W'$ und $2 \neq 0 \in \mathbb{K}$ dann treten beide Eigenwerte auf, und die zugehörigen Eigenräume sind $E_1 = W$ sowie $E_{-1} = W'$.

Unter den Eigenwerten und Eigenvektoren einer quadratischen Matrix $A \in M_{n \times n}(\mathbb{K})$ verstehen wir die Eigenwerte und Eigenvektoren der damit assoziierten linearen Abbildung $\mathbb{K}^n \rightarrow \mathbb{K}^n$, $x \mapsto Ax$. Ein Skalar $\lambda \in \mathbb{K}$ ist also genau dann Eigenwert von A , falls $0 \neq x \in \mathbb{K}^n$ existiert, sodass $Ax = \lambda x$, und jedes solche x ist Eigenvektor von A zum Eigenwert λ . Der Eigenraum zum Eigenwert λ ist gerade $E_\lambda = \{x \in \mathbb{K}^n : Ax = \lambda x\}$. Nach Proposition VI.1.3 ist λ genau dann Eigenwert von A , wenn $\det(A - \lambda I_n) = 0$ gilt. Die Matrix A wird diagonalisierbar genannt, falls die damit assoziierte lineare Abbildung $\mathbb{K}^n \rightarrow \mathbb{K}^n$, $x \mapsto Ax$, diagonalisierbar ist. Dies ist genau dann der Fall, wenn eine Basis aus Eigenvektoren b_1, \dots, b_n existiert, d.h. $Ab_i = \lambda_i b_i$. Bezeichnet $E = (e_1, \dots, e_n)$ die Standardbasis von \mathbb{K}^n , dann ist die Matrix $S := T_{EB} = (b_1 | \dots | b_n)$ invertierbar und es gilt

$$S^{-1}AS = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}, \quad (\text{VI.1})$$

denn $[\psi]_{BB} = T_{BE}[\psi]_{EE}T_{EB} = T_{EB}^{-1}AT_{EB} = S^{-1}AS$. Ist umgekehrt S eine invertierbare Matrix für die (VI.1) gilt, dann bilden die Spalten von S eine Basis aus Eigenvektoren und die Diagonaleinträge λ_i sind genau die Eigenwerte von A . Die Matrix A ist also genau dann diagonalisierbar, wenn eine invertierbare Matrix $S \in \text{GL}_n(\mathbb{K})$ existiert, sodass $S^{-1}AS$ eine Diagonalmatrix ist.

VI.1.8. DEFINITION (Ähnlichkeit von Matrizen). Zwei quadratische Matrizen $A, A' \in M_{n \times n}(\mathbb{K})$ werden *ähnlich* genannt, falls eine invertierbare Matrix $S \in \text{GL}_n(\mathbb{K})$ existiert, für die $A' = SAS^{-1}$ gilt. Wir schreiben in diesem Fall $A' \sim A$.

VI.1.9. LEMMA. *Ähnlichkeit definiert eine Äquivalenzrelation auf $M_{n \times n}(\mathbb{K})$.*

BEWEIS. Die Relation ist reflexiv, denn $A = I_n A I_n^{-1}$. Die Relation ist symmetrisch, denn aus $A' = S A S^{-1}$ folgt $A = S^{-1} A' (S^{-1})^{-1}$. Die Relation ist auch transitiv, denn aus $A' = S A S^{-1}$ und $A'' = T A' T^{-1}$ folgt $A'' = (T S) A (T S)^{-1}$. Beachte hier, dass I_n , S^{-1} und $T S$ invertierbar sind, falls dies für S und T gilt. \square

Mit dieser Terminologie ist eine quadratische Matrix A also genau dann diagonalisierbar, wenn sie ähnlich zu einer Diagonalmatrix ist. Sind $A \sim A'$ zwei ähnliche Matrizen, dann haben A und A' die gleichen Eigenwerte, und A ist genau dann diagonalisierbar, wenn A' diagonalisierbar ist. Dies folgt aus Bemerkung VI.1.5. Ist $\varphi: V \rightarrow V$ linear, und sind B, C zwei Basen von V , dann gilt $[\varphi]_{BB} \sim [\varphi]_{CC}$, denn $[\varphi]_{CC} = (T_{BC})^{-1} [\varphi]_{BB} T_{BC}$. Auch hat φ dieselben Eigenwerte wie $[\varphi]_{BB}$, es gilt daher

$$\sigma(\varphi) = \sigma([\varphi]_{BB}), \quad (\text{VI.2})$$

für jede Basis B von V . Schließlich ist die lineare Abbildung φ genau dann diagonalisierbar, wenn die Matrix $[\varphi]_{BB}$ diagonalisierbar ist, für eine (und dann jede) Basis B von V .

VI.1.10. BEISPIEL. Wir wollen Eigenwerte und Eigenvektoren der Matrix

$$A = \begin{pmatrix} 3 & 1 \\ 1 & 3 \end{pmatrix} \in M_{2 \times 2}(\mathbb{R})$$

bestimmen. Aus Proposition VI.1.3 erhalten wir folgende Gleichung für die Eigenwerte:

$$0 = \det(A - \lambda I_2) = \det \begin{pmatrix} 3 - \lambda & 1 \\ 1 & 3 - \lambda \end{pmatrix} = (3 - \lambda)^2 - 1 = (\lambda - 2)(\lambda - 4).$$

Die Matrix A hat daher genau zwei Eigenwerte, $\lambda_1 = 2$ und $\lambda_2 = 4$. Für die zugehörigen Eigenräume erhalten wir

$$\begin{aligned} E_{\lambda_1} &= \ker(A - 2I_n) = \ker \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} = \langle \begin{pmatrix} 1 \\ -1 \end{pmatrix} \rangle \\ E_{\lambda_2} &= \ker(A - 4I_n) = \ker \begin{pmatrix} -1 & 1 \\ 1 & -1 \end{pmatrix} = \langle \begin{pmatrix} 1 \\ 1 \end{pmatrix} \rangle \end{aligned}$$

Die beiden Eigenvektoren $\begin{pmatrix} 1 \\ -1 \end{pmatrix}$ und $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ bilden eine Basis von \mathbb{R}^2 , d.h. die Matrix A ist diagonalisierbar. Fassen wir diese Eigenvektoren als Spalten einer Matrix S auf, dann gilt also:

$$\underbrace{\begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}}_{S^{-1}} \underbrace{\begin{pmatrix} 3 & 1 \\ 1 & 3 \end{pmatrix}}_A \underbrace{\begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}}_S = \begin{pmatrix} 2 & 0 \\ 0 & 4 \end{pmatrix}$$

VI.1.11. BEISPIEL. Wir wollen Eigenwerte und Eigenvektoren der Matrix

$$A = \begin{pmatrix} 1 & -2 & -2 \\ 2 & 6 & 4 \\ -1 & -2 & 0 \end{pmatrix}$$

bestimmen. Da

$$\begin{aligned}\det(A - \lambda I_3) &= \det \begin{pmatrix} 1 - \lambda & -2 & -2 \\ 2 & 6 - \lambda & 4 \\ -1 & -2 & -\lambda \end{pmatrix} \\ &= (1 - \lambda)(6 - \lambda)(-\lambda) + 8 + 8 - 2(6 - \lambda) + 8(1 - \lambda) - 4\lambda \\ &= -\lambda^3 + 7\lambda^2 - 16\lambda + 12 = (2 - \lambda)^2(3 - \lambda)\end{aligned}$$

hat A genau zwei Eigenwerte, $\lambda_1 = 2$ und $\lambda_2 = 3$, siehe Proposition VI.1.3. Für die zugehörigen Eigenräume erhalten wir:

$$\begin{aligned}E_{\lambda_1} &= \ker(A - \lambda_1 I_3) = \ker \begin{pmatrix} -1 & -2 & -2 \\ 2 & 4 & 4 \\ -1 & -2 & -2 \end{pmatrix} = \left\langle \begin{pmatrix} 2 \\ -1 \\ 0 \end{pmatrix}, \begin{pmatrix} 2 \\ 0 \\ -1 \end{pmatrix} \right\rangle \\ E_{\lambda_2} &= \ker(A - \lambda_2 I_3) = \ker \begin{pmatrix} -2 & -2 & -2 \\ 2 & 3 & 4 \\ -1 & -2 & -3 \end{pmatrix} = \left\langle \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix} \right\rangle\end{aligned}$$

Der Eigenwert $\lambda_1 = 2$ hat daher geometrische Vielfachheit 2 und $\lambda_2 = 3$ hat geometrische Vielfachheit 1. Da $\begin{pmatrix} 2 \\ -1 \\ 0 \end{pmatrix}, \begin{pmatrix} 2 \\ 0 \\ -1 \end{pmatrix}, \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix}$ eine Basis aus Eigenvektoren bildet ist A diagonalisierbar und es gilt:

$$\underbrace{\begin{pmatrix} 2 & 2 & 1 \\ -1 & 0 & -2 \\ 0 & -1 & 1 \end{pmatrix}}_{S^{-1}} \underbrace{\begin{pmatrix} 1 & -2 & -2 \\ 2 & 6 & 4 \\ -1 & -2 & 0 \end{pmatrix}}_A \underbrace{\begin{pmatrix} 2 & 2 & 1 \\ -1 & 0 & -2 \\ 0 & -1 & 1 \end{pmatrix}}_S = \begin{pmatrix} 2 & & \\ & 2 & \\ & & 3 \end{pmatrix},$$

wobei die Spalten der Matrix S gerade die oben konstruierten Eigenvektoren sind.

VI.1.12. BEISPIEL. Die Eigenwerte einer Dreiecksmatrix

$$A = \begin{pmatrix} \lambda_1 & * & \cdots & * \\ 0 & \lambda_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & * \\ 0 & \cdots & 0 & \lambda_n \end{pmatrix}$$

sind genau ihre Diagonaleinträge, denn nach Satz V.1.3(n) gilt

$$\det(A - \lambda I_n) = (\lambda_1 - \lambda)(\lambda_2 - \lambda) \cdots (\lambda_n - \lambda),$$

und dies verschwindet genau dann, wenn λ mit einem der Diagonaleinträge λ_i übereinstimmt. Beachte, dass selbst eine Dreiecksmatrix nicht diagonalisierbar sein muss. Etwa hat die Matrix

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

nur einen Eigenwert, nämlich λ , und da der Eigenraum

$$E_\lambda = \ker(A - \lambda I_2) = \ker \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} = \langle \begin{pmatrix} 1 \\ 0 \end{pmatrix} \rangle$$

1-dimensional ist, existiert also keine Basis von \mathbb{K}^2 , die aus Eigenvektoren von A besteht.

VI.1.13. BEISPIEL. Für $\alpha \in \mathbb{R} \setminus \pi\mathbb{Z}$ hat die reelle Matrix

$$A = \begin{pmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{pmatrix} \in M_{2 \times 2}(\mathbb{R})$$

keinen einzigen reellen Eigenwert, denn

$$0 = \det(A - \lambda I_2) = \det \begin{pmatrix} \cos \alpha - \lambda & -\sin \alpha \\ \sin \alpha & \cos \alpha - \lambda \end{pmatrix} = (\cos \alpha - \lambda)^2 + \sin^2 \alpha$$

hat zwar die beiden Lösungen $\lambda = \cos \alpha \pm \mathbf{i} \sin \alpha$, diese sind aber nicht reell. Insbesondere ist die Matrix A über dem Körper $\mathbb{K} = \mathbb{R}$ nicht diagonalisierbar. Fassen wir A jedoch als komplexe Matrix auf,

$$A = \begin{pmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{pmatrix} \in M_{2 \times 2}(\mathbb{C}),$$

dann hat sie zwei verschiedene Eigenwerte,

$$\lambda_+ = \cos \alpha + \mathbf{i} \sin \alpha \quad \text{und} \quad \lambda_- = \cos \alpha - \mathbf{i} \sin \alpha,$$

mit entsprechenden Eigenräumen:

$$E_+ = \ker(A - \lambda_+ I_2) = \ker \begin{pmatrix} -\mathbf{i} \sin \alpha & -\sin \alpha \\ \sin \alpha & -\mathbf{i} \sin \alpha \end{pmatrix} = \langle \begin{pmatrix} 1 \\ -\mathbf{i} \end{pmatrix} \rangle$$

$$E_- = \ker(A - \lambda_- I_2) = \ker \begin{pmatrix} \mathbf{i} \sin \alpha & -\sin \alpha \\ \sin \alpha & \mathbf{i} \sin \alpha \end{pmatrix} = \langle \begin{pmatrix} 1 \\ \mathbf{i} \end{pmatrix} \rangle$$

Die beiden Vektoren $\begin{pmatrix} 1 \\ -\mathbf{i} \end{pmatrix}$ und $\begin{pmatrix} 1 \\ \mathbf{i} \end{pmatrix}$ bilden eine Basis von \mathbb{C}^2 , die aus Eigenvektoren besteht, d.h. A ist über dem Körper $\mathbb{K} = \mathbb{C}$ diagonalisierbar:

$$\underbrace{\begin{pmatrix} 1 & 1 \\ -\mathbf{i} & \mathbf{i} \end{pmatrix}^{-1}}_{S^{-1}} \underbrace{\begin{pmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{pmatrix}}_A \underbrace{\begin{pmatrix} 1 & 1 \\ -\mathbf{i} & \mathbf{i} \end{pmatrix}}_S = \begin{pmatrix} \cos \alpha + \mathbf{i} \sin \alpha & 0 \\ 0 & \cos \alpha - \mathbf{i} \sin \alpha \end{pmatrix}$$

VI.1.14. LEMMA. Sei V ein endlich-dimensionaler Vektorraum über \mathbb{K} und $\varphi: V \rightarrow V$ linear. Weiters seien $\lambda_1, \dots, \lambda_k$ paarweise verschiedene Eigenwerte von φ und $v_1, \dots, v_k \in V$ entsprechende Eigenvektoren, d.h. $v_i \neq 0$ und $\varphi(v_i) = \lambda_i v_i$, für $i = 1, \dots, k$. Dann sind die Vektoren v_1, \dots, v_k linear unabhängig in V .

BEWEIS. Wir führen den Beweis durch Induktion nach k . Der Fall $k = 1$ ist trivial, denn nach Voraussetzung gilt $v_i \neq 0$. Für den Induktionsschritt dürfen wir daher v_1, \dots, v_{k-1} als linear unabhängig voraussetzen. Seien nun $\mu_1, \dots, \mu_k \in \mathbb{K}$, sodass

$$\mu_1 v_1 + \mu_2 v_2 + \dots + \mu_{k-1} v_{k-1} + \mu_k v_k = 0. \quad (\text{VI.3})$$

Wenden wir auf diese Gleichung φ an, erhalten wir wegen $\varphi(v_i) = \lambda_i v_i$,

$$\mu_1 \lambda_1 v_1 + \mu_2 \lambda_2 v_2 + \cdots + \mu_{k-1} \lambda_{k-1} v_{k-1} + \mu_k \lambda_k v_k = 0.$$

Subtrahieren wir davon λ_k -mal die Gleichung (VI.3) folgt

$$\mu_1(\lambda_1 - \lambda_k)v_1 + \mu_2(\lambda_2 - \lambda_k)v_2 + \cdots + \mu_{k-1}(\lambda_{k-1} - \lambda_k)v_{k-1} = 0.$$

Da v_1, \dots, v_{k-1} linear unabhängig sind, muss also

$$\mu_i(\lambda_i - \lambda_k) = 0, \quad i = 1, \dots, k-1,$$

gelten. Nach Voraussetzung sind die Eigenwerte $\lambda_1, \dots, \lambda_k$ paarweise verschieden, also $\lambda_i - \lambda_k \neq 0$, für $i = 1, \dots, k-1$. Wir erhalten somit

$$\mu_1 = \mu_2 = \cdots = \mu_{k-1} = 0.$$

Mit (VI.3) folgt auch $\mu_k v_k = 0$ und da $v_k \neq 0$ schließlich $\mu_k = 0$. Dies zeigt, dass die Vektoren v_1, \dots, v_k linear unabhängig sind. \square

VI.1.15. PROPOSITION. *Sei V ein n -dimensionaler Vektorraum über \mathbb{K} und $\varphi: V \rightarrow V$ linear. Dann besitzt φ höchstens n verschiedene Eigenwerte. Hat φ genau n verschiedene Eigenwerte, so ist φ diagonalisierbar.*

BEWEIS. Dies folgt sofort aus Lemma VI.1.14. Sind $\lambda_1, \dots, \lambda_k$ paarweise verschiedene Eigenwerte von φ , dann existieren Eigenvektoren $0 \neq v_i \in V$ mit $\varphi(v_i) = \lambda_i v_i$, $i = 1, \dots, k$. Nach Lemma VI.1.14 sind die Vektoren v_1, \dots, v_k linear unabhängig, es muss daher $k \leq n$ gelten, d.h. φ kann höchstens n verschiedene Eigenwerte haben. Gilt $k = n$, dann bildet das linear unabhängige System v_1, \dots, v_n eine Basis von V , die aus Eigenvektoren von φ besteht, d.h. φ ist diagonalisierbar. \square

Nach Proposition VI.1.15 hat eine Matrix $A \in M_{n \times n}(\mathbb{K})$ also höchstens n verschiedene Eigenwerte. Sind es genau n verschiedene Eigenwerte, dann ist A diagonalisierbar. Etwa ist jede Dreiecksmatrix mit paarweise verschiedenen Diagonaleinträgen diagonalisierbar, vgl. Beispiel VI.1.12.

Proposition VI.1.15 liefert ein hinreichendes Kriterium für die Diagonalisierbarkeit linearer Abbildung und Matrizen. Für Diagonalisierbarkeit muss dieses Kriterium jedoch nicht notwendigerweise erfüllt sein. Etwa ist die identische Abbildung $\text{id}_V: V \rightarrow V$ trivialerweise diagonalisierbar, obwohl sie nur einen Eigenwert, nämlich 1 hat. Das gleiche gilt für die Einheitsmatrix I_n .

Analog zu Proposition II.5.5 haben gilt:

VI.1.16. PROPOSITION. *Sei V ein \mathbb{K} -Vektorraum und W_1, \dots, W_k Teilräume von V . Dann sind folgende Aussagen äquivalent:*

- (a) *Es gilt $V = W_1 + \cdots + W_k$ und $W_i \cap (W_1 + \cdots + W_{i-1} + W_{i+1} + \cdots + W_k) = \{0\}$, für alle $i = 1, \dots, k$.*
- (b) *Es ist $V = W_1 + \cdots + W_k$ und für alle $w_i \in W_i$ mit $w_1 + \cdots + w_k = 0$ muss schon $w_1 = \cdots = w_k = 0$ gelten.*

- (c) Zu jedem $v \in V$ existieren eindeutig bestimmte Vektoren $w_i \in W_i$, sodass $v = w_1 + \cdots + w_k$.
- (d) Ist U ein weiterer \mathbb{K} -Vektorraum und sind $\varphi_i: W_i \rightarrow U$ linear, dann existiert eine eindeutig bestimmte lineare Abbildung $\varphi: V \rightarrow U$, sodass $\varphi|_{W_i} = \varphi_i$, für alle $i = 1, \dots, k$.
- (e) Es existieren Projektoren $\pi_i: V \rightarrow V$, $\pi_i \circ \pi_i = \pi_i$, $i = 1, \dots, k$, sodass $\text{img}(\pi_i) = W_i$, $\pi_1 + \cdots + \pi_k = \text{id}_V$ und $\pi_i \circ \pi_j = 0 = \pi_j \circ \pi_i$, für alle $i \neq j$.
- Sind diese äquivalenten Eigenschaften erfüllt, und ist B_i eine Basis von W_i , $i = 1, \dots, k$, dann bildet $B_1 \cup \cdots \cup B_k$ eine Basis von V und es gilt $B_i \cap B_j = \emptyset$, für alle $i \neq j$. Im endlich-dimensionalen Fall haben wir daher die Dimensionsformel

$$\dim(V) = \dim(W_1) + \cdots + \dim(W_k).$$

BEWEIS. Ad (a) \Rightarrow (b): Schreiben wir die Gleichung $w_1 + \cdots + w_k = 0$ in der Form $-w_i = w_1 + \cdots + w_{i-1} + w_{i+1} + \cdots + w_k$, so sehen wir, dass

$$-w_i \in W_i \cap (W_1 + \cdots + W_{i-1} + W_{i+1} + \cdots + W_k), \quad i = 1, \dots, k.$$

Aus der zweiten Voraussetzung in (a) folgt daher $w_i = 0$, für alle i .

Ad (b) \Rightarrow (c): Wegen $V = W_1 + \cdots + W_k$ lässt sich jedes $v \in V$ in der Form $v = w_1 + \cdots + w_k$ schreiben, wobei $w_i \in W_i$. Um die Eindeutigkeit zu zeigen, sei $v = \tilde{w}_1 + \cdots + \tilde{w}_k$ eine weitere solche Darstellung, $\tilde{w}_i \in W_i$. Es gilt dann $0 = w'_1 + \cdots + w'_k$, wobei $w'_i := \tilde{w}_i - w_i \in W_i$. Aus der zweiten Voraussetzung in (b) folgt $w'_i = 0$, also $w_i = \tilde{w}_i$, d.h. die Darstellung $v = w_1 + \cdots + w_k$ ist eindeutig.

Ad (c) \Rightarrow (d): Da sich jedes $v \in V$ auf eindeutige Weise in der Form $v = w_1 + \cdots + w_k$ schreiben lässt, stellt

$$\varphi: V \rightarrow U, \quad \varphi(v) := \varphi_1(w_1) + \cdots + \varphi_k(w_k),$$

eine wohldefiniert Abbildung dar. Es ist noch zu zeigen, dass φ linear ist. Für $\lambda \in \mathbb{K}$ gilt $\lambda v = \lambda w_1 + \cdots + \lambda w_k$ und wegen der Linearität von φ_i daher

$$\varphi(\lambda v) = \varphi_1(\lambda w_1) + \cdots + \varphi_k(\lambda w_k) = \lambda(\varphi_1(w_1) + \cdots + \varphi_k(w_k)) = \lambda\varphi(v).$$

Ist $\tilde{v} = \tilde{w}_1 + \cdots + \tilde{w}_k$, $\tilde{w}_i \in W_i$, dann gilt $v + \tilde{v} = (w_1 + \tilde{w}_1) + \cdots + (w_k + \tilde{w}_k)$ mit $w_i + \tilde{w}_i \in W_i$ und wegen der Linearität der φ_i also

$$\begin{aligned} \varphi(v + \tilde{v}) &= \varphi_1(w_1 + \tilde{w}_1) + \cdots + \varphi_k(w_k + \tilde{w}_k) \\ &= (\varphi_1(w_1) + \cdots + \varphi_k(w_k)) + (\varphi_1(\tilde{w}_1) + \cdots + \varphi_k(\tilde{w}_k)) = \varphi(v) + \varphi(\tilde{v}). \end{aligned}$$

Damit ist die Linearität von φ gezeigt. Nach Konstruktion gilt $\varphi|_{W_i} = \varphi_i$.

Ad (d) \Rightarrow (e): Nach Voraussetzung existieren lineare Abbildung $\pi_i: V \rightarrow W_i$, $i = 1, \dots, k$, sodass $\pi_i|_{W_i} = \text{id}_{W_i}$ und $\pi_i|_{W_j} = 0$ für alle $j \neq i$. Fassen wir diese als Abbildungen $\pi_i: V \rightarrow V$ auf, dann gilt offensichtlich $\text{img}(\pi_i) = W_i$. Auch sind dies Projektoren, denn nach Konstruktion gilt $(\pi_i \circ \pi_i)|_{W_j} = \pi_i|_{W_j}$, für alle j , aus der Eindeutigkeitsaussage in (d) folgt daher $\pi_i \circ \pi_i = \pi_i$. Weiters gilt $(\pi_1 + \cdots + \pi_k)|_{W_i} = \text{id}_V|_{W_i}$, für alle i , und wegen der Eindeutigkeitsaussage in

(d) daher $\pi_1 + \dots + \pi_k = \text{id}_V$. Analog lässt sich für $i \neq j$ die Relation $\pi_i \circ \pi_j = 0$ zeigen, denn $(\pi_i \circ \pi_j)|_{W_l} = 0$, für alle $l = 1, \dots, k$.

Ad (e) \Rightarrow (a): Es ist $V = W_1 + \dots + W_k$, denn für jedes $v \in V$ haben wir $v = \text{id}_V(v) = (\pi_1 + \dots + \pi_k)(v) = \pi_1(v) + \dots + \pi_k(v)$, wobei $\pi_i(v) \in \text{img}(\pi_i) = W_i$. Aus $W_j = \text{img}(\pi_j)$ und $\pi_i \circ \pi_j = 0$ erhalten wir $\pi_i(W_j) = \{0\}$, für alle $i \neq j$. Folglich bildet π_i den gesamten Teilraum $W_1 + \dots + W_{i-1} + W_{i+1} + \dots + W_k$ auf 0 ab. Für $w \in W_i \cap (W_1 + \dots + W_{i-1} + W_{i+1} + \dots + W_k)$ gilt daher $w = \pi_i(w) = 0$, also $W_i \cap (W_1 + \dots + W_{i-1} + W_{i+1} + \dots + W_k) = \{0\}$. Damit ist die Äquivalenz der fünf Eigenschaften gezeigt.

Sei nun B_i eine Basis von W_i , $i = 1, \dots, k$. Für $i \neq j$ gilt $B_i \cap B_j = \emptyset$, denn $B_i \cap B_j \subseteq W_i \cap W_j \subseteq W_i \cap (W_1 + \dots + W_{i-1} + W_{i+1} + \dots + W_k) = \{0\}$. Da $V = W_1 + \dots + W_k$, ist $B := B_1 \cup \dots \cup B_k$ ein Erzeugendensystem von V . Sei nun $\sum_{b \in B} \lambda_b b = 0$, wobei fast alle $\lambda_b \in \mathbb{K}$ verschwinden. Dann gilt $\sum_{i=1}^k \sum_{b \in B_i} \lambda_b b = 0$, und wegen (b) daher $\sum_{b \in B_i} \lambda_b b = 0$, für jedes $i = 1, \dots, k$. Aufgrund der linearen Unabhängigkeit von B_i erhalten wir $\lambda_b = 0$, für alle $b \in B$. Dies zeigt, dass B linear unabhängig, also eine Basis von V ist. \square

VI.1.17. DEFINITION (Direkte Summe). Sind die äquivalenten Eigenschaften in Proposition VI.1.16 erfüllt, dann schreiben wir

$$V = \bigoplus_{i=1}^k W_i \quad \text{oder} \quad V = W_1 \oplus \dots \oplus W_k, \quad (\text{VI.4})$$

und sagen V ist die *direkte Summe* der Teilräume W_1, \dots, W_k . Offensichtlich ist dies eine Verallgemeinerung der in Definition II.5.6 eingeführten direkten Summe zweier Teilräume. Die Projektoren aus Proposition VI.1.16(e) werden als die mit der Zerlegung (VI.4) *assozierten Projektionen* oder als die damit assoziierte *Zerlegung der Eins* bezeichnet.

VI.1.18. SATZ (Diagonalisierbarkeit). Sei V ein n -dimensionaler Vektorraum über \mathbb{K} , und $\varphi: V \rightarrow V$ linear. Weiters bezeichnen $\lambda_1, \dots, \lambda_k$ alle (verschiedenen) Eigenwerte und $E_{\lambda_i} = \ker(\varphi - \lambda_i \text{id}_V)$ die zugehörigen Eigenräume von φ . Dann sind folgende Aussagen äquivalent:

- (a) φ ist diagonalisierbar.
- (b) Es gilt $\dim(E_{\lambda_1}) + \dots + \dim(E_{\lambda_k}) = n$, d.h. die Summe der geometrischen Vielfachheiten aller Eigenwerte stimmt mit $\dim(V)$ überein.
- (c) Es gilt $V = E_{\lambda_1} \oplus \dots \oplus E_{\lambda_k}$.

Ist φ diagonalisierbar, und bezeichnen $\pi_i: V \rightarrow V$ die mit der Zerlegung $V = E_{\lambda_1} \oplus \dots \oplus E_{\lambda_k}$ assoziierten Projektoren, dann gilt (Spektralzerlegung von φ)

$$\varphi = \lambda_1 \pi_1 + \dots + \lambda_k \pi_k. \quad (\text{VI.5})$$

Weiters existiert eine Basis B von V , sodass

$$[\varphi]_{BB} = \begin{pmatrix} \lambda_1 I_{m_1} & & \\ & \ddots & \\ & & \lambda_k I_{m_k} \end{pmatrix}, \quad (\text{VI.6})$$

wobei $m_i = \dim(E_{\lambda_i})$ die geometrischen Vielfachheiten bezeichnen.

BEWEIS. Es bezeichne $E = E_{\lambda_1} + \cdots + E_{\lambda_k}$ die Summe aller Eigenräume. Dies ist eine direkte Summe, d.h. es gilt

$$E = E_{\lambda_1} \oplus \cdots \oplus E_{\lambda_k}. \quad (\text{VI.7})$$

Sind nämlich $v_i \in E_{\lambda_i}$ und $v_1 + \cdots + v_k = 0$, dann folgt mit Lemma VI.1.14 sofort $v_1 = \cdots = v_k = 0$, siehe auch Proposition VI.1.16(b). Insbesondere haben wir

$$\dim(E) = \dim(E_{\lambda_1}) + \cdots + \dim(E_{\lambda_k}). \quad (\text{VI.8})$$

Ad (a) \Rightarrow (b): Nach Voraussetzung existiert eine Basis b_1, \dots, b_n von V , die aus Eigenvektoren besteht. Da jeder dieser Basisvektoren in einem der Eigenräume E_{λ_i} liegt, enthält E eine linear unabhängige Teilmenge mit n Elementen, es ist daher $\dim(E) \geq n$. Wegen $E \subseteq V$ muss aber auch $\dim(E) \leq n$ gelten. Somit ist $\dim(E) = n$. Aus (VI.8) erhalten wir nun $\dim(E_{\lambda_1}) + \cdots + \dim(E_{\lambda_k}) = n$.

Ad (b) \Rightarrow (c): Nach Voraussetzung gilt $\dim(V) = \dim(E_{\lambda_1}) + \cdots + \dim(E_{\lambda_k})$. Zusammen mit (VI.8) folgt $\dim(E) = \dim(V)$ und daher $E = V$, da ja $E \subseteq V$. Aus (VI.7) folgt somit $V = E_{\lambda_1} \oplus \cdots \oplus E_{\lambda_k}$.

Ad (c) \Rightarrow (a): Für $i = 1, \dots, k$ sei $b_1^{(i)}, \dots, b_{m_i}^{(i)}$ eine Basis des Eigenraums E_{λ_i} , wobei $m_i = \dim(E_{\lambda_i})$. Nach Proposition VI.1.16 ist dann

$$b_1^{(1)}, \dots, b_{m_1}^{(1)}, b_1^{(2)}, \dots, b_{m_2}^{(2)}, \dots, b_1^{(k)}, \dots, b_{m_k}^{(k)}$$

eine Basis von $V = E_{\lambda_1} \oplus \cdots \oplus E_{\lambda_k}$. Nach Konstruktion besteht diese Basis aus Eigenvektoren von φ , also ist φ diagonalisierbar. Die Matrixdarstellung von φ bezüglich dieser Basis hat die Form (VI.6). Die Gleichung (VI.5) folgt aus der Eindeutigkeitsaussage in Proposition VI.1.16(d), denn offensichtlich gilt $\varphi|_{E_{\lambda_i}} = \lambda_i \text{id}_{E_{\lambda_i}} = (\lambda_1 \pi_1 + \cdots + \lambda_k \pi_k)|_{E_{\lambda_i}}$, für alle $i = 1, \dots, k$. \square

Ist $p \in \mathbb{K}[z]$ ein Polynom, $p = \sum_i p_i z^i$, und $A \in M_{n \times n}(\mathbb{K})$ eine quadratische Matrix, dann können wir A in p einsetzen und erhalten eine Matrix $p(A) \in M_{n \times n}(\mathbb{K})$, genauer

$$p(A) := \sum_i p_i A^i = p_0 I_n + p_1 A + p_2 A^2 + p_3 A^3 + \cdots$$

Analog definieren wir für jeden Endomorphismus $\varphi: V \rightarrow V$,

$$p(\varphi) := \sum_i p_i \varphi^i = p_0 \text{id}_V + p_1 \varphi + p_2 \varphi^2 + p_3 \varphi^3 + \cdots$$

wobei $\varphi^i = \varphi \circ \cdots \circ \varphi$ und $\varphi^0 = \text{id}_V$.

VI.1.19. LEMMA. Sind $p, q \in \mathbb{K}[z]$, $\lambda \in \mathbb{K}$ und $A \in M_{n \times n}(\mathbb{K})$, dann gilt

$$(p + q)(A) = p(A) + q(A), \quad (\lambda p)(A) = \lambda p(A) \quad \text{und} \quad (pq)(A) = p(A)q(A),$$

d.h. $\mathbb{K}[z] \rightarrow M_{n \times n}(\mathbb{K})$, $p \mapsto p(A)$, ist ein Homomorphismus von \mathbb{K} -Algebren. Weiters gilt $p(A^t) = p(A)^t$ und $p(SAS^{-1}) = Sp(A)S^{-1}$, für jedes $S \in \text{GL}_n(\mathbb{K})$.

Ist V ein \mathbb{K} -Vektorraum und $\varphi \in \text{end}(V)$, dann haben wir analog

$$(p + q)(\varphi) = p(\varphi) + q(\varphi), \quad (\lambda p)(\varphi) = \lambda p(\varphi) \quad \text{und} \quad (pq)(\varphi) = p(\varphi)q(\varphi),$$

d.h. $\mathbb{K}[z] \rightarrow \text{end}(V)$, $p \mapsto p(\varphi)$, ist ein Homomorphismus von \mathbb{K} -Algebren. Weiters gilt $p(\varphi^t) = p(\varphi)^t$ und $p(\psi \circ \varphi \circ \psi^{-1}) = \psi \circ p(\varphi) \circ \psi^{-1}$, für jeden linearen Isomorphismus $\psi: V \xrightarrow{\cong} W$. Ist V endlich-dimensional und B eine geordnete Basis von V , dann gilt $[p(\varphi)]_{BB} = p([\varphi]_{BB})$.

BEWEIS. Bezeichnen $p = \sum_i p_i z^i$ und $q = \sum_i q_i z^i$ zwei Polynome, dann folgt

$$\begin{aligned} (p + q)(A) &= (\sum_i p_i z^i + \sum_i q_i z^i)(A) = (\sum_i (p_i + q_i) z^i)(A) \\ &= \sum_i (p_i + q_i) A^i = \sum_i p_i A^i + \sum_i q_i A^i = p(A) + q(A). \end{aligned}$$

Für $\lambda \in \mathbb{K}$ erhalten wir analog

$$(\lambda p)(A) = (\lambda \sum_i p_i z^i)(A) = (\sum_i (\lambda p_i) z^i)(A) = \sum_i (\lambda p_i) A^i = \lambda \sum_i p_i A^i = \lambda p(A).$$

Ebenso

$$\begin{aligned} (pq)(A) &= \left((\sum_i p_i z^i) (\sum_j q_j z^j) \right) (A) \\ &= \left(\sum_k (\sum_{i+j=k} p_i q_j) z^k \right) (A) = \sum_k (\sum_{i+j=k} p_i q_j) A^k \\ &= \sum_k \sum_{i+j=k} p_i A^i q_j A^j = (\sum_i p_i A^i) (\sum_j q_j A^j) = p(A)q(A) \end{aligned}$$

und $p(A^t) = (\sum_i p_i z^i)(A^t) = \sum_i p_i (A^t)^i = \sum_i p_i (A^i)^t = (\sum_i p_i A^i)^t = p(A)^t$. Für $S \in \text{GL}_n(\mathbb{K})$ gilt $(SAS^{-1})^i = (SAS^{-1}) \cdots (SAS^{-1}) = SA^i S^{-1}$ und daher

$$p(SAS^{-1}) = \sum_i p_i (SAS^{-1})^i = \sum_i p_i SA^i S^{-1} = S(\sum_i p_i A^i) S^{-1} = Sp(A)S^{-1}.$$

Die entsprechenden Eigenschaften von $p(\varphi)$ lassen sich völlig analog herleiten. Ist schließlich V endlich-dimensional und B eine geordnete Basis von V , dann gilt

$$[p(\varphi)]_{BB} = [\sum_i p_i \varphi^i]_{BB} = \sum_i p_i [\varphi^i]_{BB} = \sum_i p_i ([\varphi]_{BB})^i = p([\varphi]_{BB}),$$

wobei wir die Eigenschaften in Satz IV.6.15 verwendet haben. \square

VI.1.20. BEISPIEL. Für $p = 2 + 3z + 5z^2$ und $A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$ ist

$$p(A) = 2I_2 + 3A + 5A^2 = 2 \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + 3 \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} + 5 \begin{pmatrix} 7 & 10 \\ 15 & 22 \end{pmatrix} = \begin{pmatrix} 40 & 56 \\ 84 & 124 \end{pmatrix}.$$

VI.1.21. BEISPIEL. Ist $p \in \mathbb{K}[z]$ ein Polynom und D eine Dreiecksmatrix,

$$D = \begin{pmatrix} \lambda_1 & * & \cdots & * \\ & \lambda_2 & \ddots & \vdots \\ & & \ddots & * \\ & & & \lambda_n \end{pmatrix}, \quad \text{dann gilt} \quad p(D) = \begin{pmatrix} p(\lambda_1) & * & \cdots & * \\ & p(\lambda_2) & \ddots & \vdots \\ & & \ddots & * \\ & & & p(\lambda_n) \end{pmatrix},$$

siehe Aufgabe 35. Für die Spektra gilt daher $\sigma(p(D)) = p(\sigma(D))$.

VI.1.22. BEMERKUNG. Beachte, dass i.A. $p(\lambda A) \neq \lambda p(A)$, $p(A+B) \neq p(A) + p(B)$ und $p(AB) \neq p(A)p(B)$ gilt, wobei $\lambda \in \mathbb{K}$, $p \in \mathbb{K}[z]$ und $A, B \in M_{n \times n}(\mathbb{K})$.

VI.1.23. PROPOSITION. Sei V ein endlich-dimensionaler Vektorraum über \mathbb{K} , $\varphi: V \rightarrow V$ eine lineare Abbildung und $p \in \mathbb{K}[z]$ ein Polynom. Ist v Eigenvektor von φ zum Eigenwert λ , dann ist v auch Eigenvektor von $p(\varphi)$ zum Eigenwert $p(\lambda)$. Für die Spektra gilt daher $p(\sigma(\varphi)) \subseteq \sigma(p(\varphi))$. Existiert eine Basis B von V , sodass $[\varphi]_{BB}$ obere Dreiecksgestalt hat, dann haben wir sogar $p(\sigma(\varphi)) = \sigma(p(\varphi))$.

BEWEIS. Ist v Eigenvektor zum Eigenwert λ , d.h. $\varphi(v) = \lambda v$, dann folgt $p(\varphi)(v) = (\sum_i p_i \varphi^i)(v) = \sum_i p_i \varphi^i(v) = \sum_i p_i \lambda^i v = p(\lambda)v$, also ist v auch Eigenvektor von $p(\varphi)$ zum Eigenwert $p(\lambda)$. Insbesondere gilt $p(\sigma(\varphi)) \subseteq \sigma(p(\varphi))$. Sei nun B eine Basis von V , sodass $[\varphi]_{BB}$ obere Dreiecksgestalt hat. Aus Beispiel VI.1.21 folgt $p(\sigma([\varphi]_{BB})) = \sigma(p([\varphi]_{BB}))$ und somit

$$p(\sigma(\varphi)) = p(\sigma([\varphi]_{BB})) = \sigma(p([\varphi]_{BB})) = \sigma([p(\varphi)]_{BB}) = \sigma(p(\varphi)),$$

wobei wir auch zwei mal (VI.2) verwendet haben. \square

Für eine Matrix $A \in M_{n \times n}(\mathbb{K})$ und ein Polynom $p \in \mathbb{K}[z]$ besagt Proposition VI.1.23, dass jeder Eigenvektor von A zum Eigenwert λ auch ein Eigenvektor von $p(A)$ ist, und zwar zum Eigenwert $p(\lambda)$. Für die Spektra gilt daher $p(\sigma(A)) \subseteq \sigma(p(A))$. Ist A ähnlich zu einer oberen Dreiecksmatrix, dann haben wir sogar $p(\sigma(A)) = \sigma(p(A))$.

VI.1.24. BEISPIEL. Betrachte das Polynom $p = z^2$ und die reelle Matrix $A = \begin{pmatrix} 1 & \\ -1 & 2 \end{pmatrix}$. Dann gilt $p(A) = \begin{pmatrix} 1 & \\ 1 & 4 \end{pmatrix}$, also $p(\sigma(A)) = \sigma(p(A))$. Beachte jedoch, dass A drei 1-dimensionale Eigenräume besitzt, während $p(A)$ nur zwei Eigenräume hat von denen einer 2-dimensional ist.

VI.1.25. BEISPIEL. In Beispiel VI.1.13 haben wir gesehen, dass die reelle Matrix $A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ über $\mathbb{K} = \mathbb{R}$ keinen Eigenwert besitzt, d.h. $\sigma(A) = \emptyset$. Setzen wir A in das Polynom $p = z^2$ ein, erhalten wir $p(A) = A^2 = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$, eine diagonalisierbare Abbildung mit Spektrum $\sigma(p(A)) = \{-1\}$. Dieses Beispiel zeigt, dass i.A. $p(\sigma(A)) \neq \sigma(p(A))$.

Wir beenden diesen Abschnitt mit einer ersten Anwendung:

VI.1.26. BEISPIEL (Fibonacci-Folge). Betrachte die rekursiv definierte Fibonacci-Folge: $0, 1, 1, 2, 3, 5, 8, 13, 21, 34, \dots$. Genauer:

$$x_0 = 0, \quad x_1 = 1, \quad x_n = x_{n-1} + x_{n-2}, \quad n \geq 2.$$

Wir wollen eine geschlossene Formel für x_n herleiten. Dazu betrachten wir die Vektoren $v_n := \begin{pmatrix} x_n \\ x_{n+1} \end{pmatrix} \in \mathbb{R}^2$ und die Matrix $A = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}$. Es gilt dann

$$v_0 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \quad v_n = Av_{n-1}, \quad n \geq 1.$$

Somit ist $v_n = A^n v_0$ und x_n also die erste Komponente von $A^n v_0$. Es genügt daher eine geschlossene Formel für A^n herzuleiten. Dies lässt sich durch Diagonalisieren von A erreichen. Lösen der quadratischen Gleichung $0 = \det(A - \lambda I_2) = \det \begin{pmatrix} -\lambda & 1 \\ 1 & 1-\lambda \end{pmatrix} = \lambda^2 - \lambda - 1$ liefert die beiden verschiedenen Eigenwerte

$$\lambda_{\pm} = \frac{1 \pm \sqrt{5}}{2}.$$

Die Matrix A ist also diagonalisierbar, d.h. es existiert eine invertierbare Matrix S , sodass $S^{-1}AS = D$, wobei $D = \begin{pmatrix} \lambda_+ & 0 \\ 0 & \lambda_- \end{pmatrix}$. Für die Potenzen von A folgt

$$A^n = (SDS^{-1})^n = (SDS^{-1}) \cdots (SDS^{-1}) = SD^n S^{-1} = S \begin{pmatrix} \lambda_+^n & 0 \\ 0 & \lambda_-^n \end{pmatrix} S^{-1}.$$

Um eine Matrix S wie oben zu bestimmen, betrachten wir die beiden Eigenräume:

$$E_+ = \ker(A - \lambda_+ I_2) = \ker \begin{pmatrix} -\frac{1+\sqrt{5}}{2} & 1 \\ 1 & \frac{1-\sqrt{5}}{2} \end{pmatrix} = \left\langle \begin{pmatrix} 1 \\ \frac{1+\sqrt{5}}{2} \end{pmatrix} \right\rangle$$

$$E_- = \ker(A - \lambda_- I_2) = \ker \begin{pmatrix} -\frac{1-\sqrt{5}}{2} & 1 \\ 1 & \frac{1+\sqrt{5}}{2} \end{pmatrix} = \left\langle \begin{pmatrix} 1 \\ \frac{1-\sqrt{5}}{2} \end{pmatrix} \right\rangle$$

Somit

$$S = \begin{pmatrix} 1 & 1 \\ \frac{1+\sqrt{5}}{2} & \frac{1-\sqrt{5}}{2} \end{pmatrix} \quad \text{und} \quad S^{-1} = \frac{1}{\sqrt{5}} \begin{pmatrix} -\frac{1-\sqrt{5}}{2} & 1 \\ \frac{1+\sqrt{5}}{2} & -1 \end{pmatrix}.$$

Es folgt

$$\begin{pmatrix} x_n \\ x_{n+1} \end{pmatrix} = v_n = A^n v_0 = S \begin{pmatrix} \lambda_+^n & 0 \\ 0 & \lambda_-^n \end{pmatrix} S^{-1} \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

$$= \frac{1}{\sqrt{5}} S \begin{pmatrix} \lambda_+^n & 0 \\ 0 & \lambda_-^n \end{pmatrix} \begin{pmatrix} 1 \\ -1 \end{pmatrix} = \frac{1}{\sqrt{5}} S \begin{pmatrix} \lambda_+^n \\ -\lambda_-^n \end{pmatrix} = \frac{1}{\sqrt{5}} \begin{pmatrix} \lambda_+^n - \lambda_-^n \\ \lambda_+^{n+1} - \lambda_-^{n+1} \end{pmatrix}$$

also $x_n = \frac{\lambda_+^n - \lambda_-^n}{\sqrt{5}}$ und daher

$$x_n = \frac{\left(\frac{1+\sqrt{5}}{2}\right)^n - \left(\frac{1-\sqrt{5}}{2}\right)^n}{\sqrt{5}},$$

die gesuchte geschlossene Formel für die n -te Fibonacci-Zahl. Mit dieser Methode lassen sich natürlich viel allgemeinere lineare Rekursionen lösen.

VI.2. Charakteristisches Polynom. Die die Eigenwerte charakterisierende Gleichung $\det(\varphi - \lambda \text{id}_V) = 0$, siehe Proposition VI.1.3, ist polynomial in λ . Wir wollen dieses Polynom nun genauer untersuchen und damit u.a. eine weitere Charakterisierung der Diagonalisierbarkeit herleiten.

Unter dem *Grad* eines Polynoms $p = p_0 + p_1z + p_2z^2 + \dots$ verstehen wir die größte Zahl n , für die $p_n \neq 0$ gilt. Wir schreiben dafür $\deg(p) = n$. Der Grad des Nullpolynoms wird durch $\deg(0) := -\infty$ definiert. Mit dieser Konvention gilt

$$\deg(pq) = \deg(p) + \deg(q),$$

für beliebige Polynome p und q . Beachte, dass ein Polynom genau dann konstant ist, wenn sein Grad kleiner oder gleich Null ist. Polynome vom Grad 1 sind genau die Polynome der Form $p = p_0 + p_1z$, wobei $p_0, p_1 \in \mathbb{K}$ und $p_1 \neq 0$.

VI.2.1. LEMMA (Polynomdivision). *Seien $p, q \in \mathbb{K}[z]$ Polynome und $q \neq 0$. Dann existieren eindeutig bestimmte Polynome $s, r \in \mathbb{K}[z]$, sodass*

$$p = qs + r \quad \text{und} \quad \deg(r) < \deg(q).$$

BEWEIS. Wir beginnen mit der Eindeutigkeit der Darstellung. Seien dazu $s, \tilde{s}, r, \tilde{r} \in \mathbb{K}[z]$, sodass $qs + r = p = q\tilde{s} + \tilde{r}$, $\deg(r) < \deg(q)$ und $\deg(\tilde{r}) < \deg(q)$. Es folgt $q(s - \tilde{s}) = \tilde{r} - r$ und $\deg(\tilde{r} - r) < \deg(q)$. Somit $\deg(q(s - \tilde{s})) < \deg(q)$, also $s - \tilde{s} = 0$. Dies zeigt, $s = \tilde{s}$, woraus aber auch sofort $r = \tilde{r}$ folgt. Damit ist die Eindeutigkeit der Darstellung gezeigt.

Sei $q = \sum_{i=0}^m q_i z^i$, wobei $m = \deg(q) \geq 0$ und daher $q_m \neq 0$. Ist $\deg(p) < m$, dann haben $s = 0$ und $r = p$ die gewünschten Eigenschaften. O.B.d.A. sei daher $n := \deg(p) \geq m$. Wir führen den Beweis mittels Induktion nach dem Grad von p . Bezeichne dazu $p = \sum_{i=0}^n p_i z^i$, also $p_n \neq 0$. Betrachten wir das Polynom $\tilde{s} := \frac{p_n}{q_m} z^{n-m}$ dann gilt $\deg(p - q\tilde{s}) < \deg(p)$. Nach Induktionsvoraussetzung existieren daher Polynome \bar{s} und r , sodass $p - q\tilde{s} = q\bar{s} + r$ und $\deg(r) < \deg(q)$. Setzen wir $s := \tilde{s} + \bar{s}$, so folgt $p = qs + r$, wie gewünscht. \square

VI.2.2. BEISPIEL. Mit dem bekannten Algorithmus erhalten wir etwa:

$$\underbrace{x^5 - 2x^4 + 2x^3 + 9x^2 - 12x + 10}_p = \underbrace{(x^2 - 2x + 3)}_q \underbrace{(x^3 - x + 7)}_s + \underbrace{5x - 11}_r.$$

Beachte, dass dieser Algorithmus auf der im Beweis von Lemma VI.2.1 verwendeten Konstruktion basiert.

VI.2.3. LEMMA. *Sei $p \in \mathbb{K}[z]$ ein Polynom und $\lambda \in \mathbb{K}$ eine Nullstelle von p . Dann existiert ein eindeutig bestimmtes Polynom $q \in \mathbb{K}[z]$, sodass $p = (z - \lambda)q$.*

BEWEIS. Nach Lemma VI.2.1 existieren eindeutig bestimmte Polynome q und r , sodass $p = (z - \lambda)q + r$ und $\deg(r) < \deg(z - \lambda) = 1$. Somit ist r ein konstantes Polynom, d.h. $r = r_0$, wobei $r_0 \in \mathbb{K}$. Einsetzen von λ liefert

$$0 = p(\lambda) = (\lambda - \lambda)q(\lambda) + r(\lambda) = r_0,$$

also $r = 0$, und daher $p = (z - \lambda)q$. \square

VI.2.4. LEMMA. *Ist $0 \neq p \in \mathbb{K}[z]$ ein nicht-triviales Polynom, dann existieren $\lambda_1, \dots, \lambda_k \in \mathbb{K}$ und ein Polynom $q \in \mathbb{K}[z]$, ohne Nullstellen in \mathbb{K} , sodass*

$$p = (z - \lambda_1) \cdots (z - \lambda_k)q.$$

BEWEIS. Wir führen den Beweis mittels Induktion nach dem Grad des Polynoms p . Der Induktionsanfang, $\deg(p) = 0$, ist trivial. Nun zum Induktionsschritt: Besitzt p keine Nullstelle in \mathbb{K} , dann haben $k = 0$ und $q = p$ die gewünschten Eigenschaften. O.B.d.A. sei daher $\lambda_1 \in \mathbb{K}$ eine Nullstelle von p . Nach Lemma VI.2.3 existiert ein Polynom $\tilde{p} \in \mathbb{K}[z]$, sodass $p = (z - \lambda_1)\tilde{p}$. Beachte $\tilde{p} \neq 0$ und $\deg(\tilde{p}) = \deg(p) - 1$. Nach Induktionsvoraussetzung existieren daher $\lambda_2, \dots, \lambda_k \in \mathbb{K}$ und ein Nullstellen-freies Polynom $q \in \mathbb{K}[z]$, sodass

$$\tilde{p} = (z - \lambda_2) \cdots (z - \lambda_k)q.$$

Insgesamt folgt $p = (z - \lambda_1)\tilde{p} = (z - \lambda_1)(z - \lambda_2) \cdots (z - \lambda_k)q$. Damit ist der Induktionsschritt gezeigt und das Lemma bewiesen. \square

VI.2.5. PROPOSITION. *Sei \mathbb{K} ein Körper und $0 \neq p \in \mathbb{K}[z]$ ein Polynom vom Grad n . Dann hat p höchstens n verschiedene Nullstellen in \mathbb{K} . Bezeichnen $\lambda_1, \dots, \lambda_k \in \mathbb{K}$ alle, paarweise verschiedenen, Nullstellen von p , dann existieren eindeutig bestimmte Zahlen $n_1, \dots, n_k \in \mathbb{N}$ und ein eindeutig bestimmtes Polynom $q \in \mathbb{K}[z]$ ohne Nullstellen in \mathbb{K} , sodass*

$$p = (z - \lambda_1)^{n_1} \cdots (z - \lambda_k)^{n_k} q. \quad (\text{VI.9})$$

Weiters gilt $n = n_1 + \cdots + n_k + \deg(q)$ und daher auch $n_1 + \cdots + n_k \leq n$.

BEWEIS. Durch Zusammenfassen mehrfach auftretender Linearfaktoren in Lemma VI.2.4 erhalten wir $n_1, \dots, n_k \in \mathbb{N}$, ein Nullstellen-freies Polynom $q \in \mathbb{K}[z]$ und paarweise verschiedene $\lambda_1, \dots, \lambda_k \in \mathbb{K}$, sodass

$$p = (z - \lambda_1)^{n_1} \cdots (z - \lambda_k)^{n_k} q.$$

Offensichtlich ist jedes λ_i eine Nullstelle von p . Andererseits sind dies schon alle Nullstellen von p , denn aus $p(\lambda) = 0$ folgt $(\lambda - \lambda_1)^{n_1} \cdots (\lambda - \lambda_k)^{n_k} q(\lambda) = 0$, also $\lambda = \lambda_i$ für ein i , da ja $q(\lambda) \neq 0$. Für den Grad von p erhalten wir

$$n = \deg(p) = n_1 + \cdots + n_k + \deg(q) \geq n_1 + \cdots + n_k.$$

Inbesondere folgt $k \leq n$, also kann p höchstens n verschiedene Nullstellen haben.

Um die Eindeutigkeit von n_i und q zu beweisen werden wir folgende Kürzungsregel verwenden: Ist $sr = s\tilde{r}$, wobei $s, r, \tilde{r} \in \mathbb{K}[z]$ und $s \neq 0$, dann gilt schon $r = \tilde{r}$. Dies folgt sofort aus der Eindeutigkeitsaussage in Lemma VI.2.1.

Sei nun $p = (z - \lambda_1)^{\tilde{n}_1} \cdots (z - \lambda_k)^{\tilde{n}_k} \tilde{q}$ eine weitere solche Darstellung, d.h. $\tilde{n}_1, \dots, \tilde{n}_k \in \mathbb{N}$ und $\tilde{q} \in \mathbb{K}[z]$ Nullstellen-frei. O.B.d.A. sei $\tilde{n}_1 \geq n_1$. Wenden wir oben erwähnte Kürzungsregel mit $s = (z - \lambda_1)^{n_1}$ auf

$$(z - \lambda_1)^{n_1} \cdots (z - \lambda_k)^{n_k} q = p = (z - \lambda_1)^{\tilde{n}_1} \cdots (z - \lambda_k)^{\tilde{n}_k} \tilde{q}$$

an, so erhalten wir

$$(z - \lambda_2)^{n_2} \cdots (z - \lambda_k)^{n_k} q = (z - \lambda_1)^{\tilde{n}_1 - n_1} (z - \lambda_2)^{\tilde{n}_2} \cdots (z - \lambda_k)^{\tilde{n}_k} \tilde{q}.$$

Da die linke Seite für $z = \lambda_1$ nicht verschwindet, muss dies auch für die rechte Seite gelten. Daraus folgt $\tilde{n}_1 = n_1$ und dann

$$(z - \lambda_2)^{n_2} \cdots (z - \lambda_k)^{n_k} q = (z - \lambda_2)^{\tilde{n}_2} \cdots (z - \lambda_k)^{\tilde{n}_k} \tilde{q}.$$

Induktiv fortfahrend, erhalten wir $n_i = \tilde{n}_i$, für jedes i und schließlich $q = \tilde{q}$. \square

VI.2.6. DEFINITION (Algebraische Vielfachheit). Sei $0 \neq p \in \mathbb{K}[z]$ ein Polynom vom Grad n und $p = (z - \lambda_1)^{n_1} \cdots (z - \lambda_k)^{n_k} q$ wie in Proposition VI.2.5. Dann wird n_i die *algebraische Vielfachheit* der Nullstelle λ_i genannt. Beachte, dass dies nach der Eindeutigkeitsaussage in Proposition VI.2.5 wohldefiniert ist.

Sei $0 \neq p \in \mathbb{K}[z]$ ein Polynom vom Grad n . Nach Proposition VI.2.5 ist die Summe der algebraischen Vielfachheiten aller Nullstellen von p höchstens n . Dies wird oft auch so formuliert: p hat höchstens n Nullstellen, wenn diese entsprechend ihrer Vielfachheit gezählt werden. Wir sagen das Polynom p *zerfällt (über \mathbb{K}) in Linearfaktoren*, wenn es sich in der Form

$$p = c(z - \lambda_1) \cdots (z - \lambda_n) \tag{VI.10}$$

schreiben lässt, wobei $\lambda_1, \dots, \lambda_n \in \mathbb{K}$ und $c \in \mathbb{K}$. Dies ist also genau dann der Fall, wenn das Polynom q in Proposition VI.2.5 konstant ist, d.h. wenn die Summe der algebraischen Vielfachheiten aller Nullstellen gleich n ist. Ein Polynom vom Grad n zerfällt daher in Linearfaktoren, wenn es genau n Nullstellen (mit Vielfachheit gezählt) besitzt. In diesem Fall ist die Darstellung (VI.10) bis auf Permutation der Linearfaktoren eindeutig, und $\lambda_1, \dots, \lambda_n$ sind alle Nullstellen von p , wobei diese ihrer algebraischen Vielfachheit entsprechend oft auftreten.

VI.2.7. BEISPIEL. Betrachte das reelle Polynom $p \in \mathbb{R}[z]$,

$$p = z^6 - 5z^5 + 10z^4 - 12z^3 + 11z^2 - 7z + 2.$$

Durch Erraten von Nullstellen und Polynomdivision erhalten wir

$$p = (z - 1)^3(z - 2)(z^2 + 1).$$

Das Polynom p hat daher zwei reelle Nullstellen, $\lambda_1 = 1$ mit Vielfachheit 3 und $\lambda_2 = 2$ mit Vielfachheit 1. Es zerfällt über \mathbb{R} nicht in Linearfaktoren, denn $z^2 + 1$ besitzt keine reelle Nullstelle. Fassen wir p als komplexes Polynom auf, $p \in \mathbb{C}[z]$, dann gilt

$$p = (z - 1)^3(z - 2)(z - \mathbf{i})(z + \mathbf{i}).$$

Über \mathbb{C} zerfällt p daher in Linearfaktoren. Zu den bereits genannten reellen Nullstellen kommen nun noch zwei komplexe Nullstellen, $\lambda_3 = \mathbf{i}$ und $\lambda_4 = -\mathbf{i}$, jeweils mit Vielfachheit 1, hinzu.

VI.2.8. DEFINITION (Algebraisch abgeschlossene Körper). Ein Körper \mathbb{K} wird *algebraisch abgeschlossen* genannt, falls jedes nicht konstante Polynom $p \in \mathbb{K}[z]$ eine Nullstelle $\lambda \in \mathbb{K}$ besitzt, $p(\lambda) = 0$.

Aus Proposition VI.2.5 erhalten wir sofort:

VI.2.9. PROPOSITION. *Über einem algebraisch abgeschlossenen Körper zerfällt jedes nicht-triviale Polynom vom Grad n in Linearfaktoren und besitzt genau n Nullstellen, wenn diese ihrer Vielfachheit entsprechend oft gezählt werden.*

Der Körper \mathbb{Q} ist nicht algebraisch abgeschlossen, denn das Polynom $p = z^2 - 2$ besitzt keine Nullstelle in \mathbb{Q} . Auch der Körper \mathbb{R} ist nicht algebraisch abgeschlossen, denn das Polynom $p = z^2 + 1$ besitzt keine reelle Nullstelle.

VI.2.10. SATZ (Fundamentalsatz der Algebra). *Der Körper \mathbb{C} ist algebraisch abgeschlossen.*

BEWEIS. Sei also $p = p_0 + p_1z + \cdots + p_nz^n$ ein komplexes Polynom, $p_i \in \mathbb{C}$, $n \geq 1$ und $p_n \neq 0$. Es ist zu zeigen, dass $z \in \mathbb{C}$ existiert, für das $p(z) = 0$ gilt.

Wir leiten zunächst folgende Abschätzung her: Es existieren reelle Zahlen $0 < m \leq M$ und $R \geq 0$, sodass

$$m|z|^n \leq |p(z)| \leq M|z|^n, \quad \text{für alle } z \in \mathbb{C} \text{ mit } |z| \geq R. \quad (\text{VI.11})$$

Wir zeige, dass $m := |p_n|/2$, $M := \sum_{i=0}^n |p_i|$ und $R := \max\{1, 2M/|p_n|\}$ die gewünschte Eigenschaft haben. Sei dazu $z \in \mathbb{C}$ mit $|z| \geq R$. Insbesondere gilt daher $|z| \geq 1$. Mit der Dreiecksungleichung erhalten wir sofort:

$$|p(z)| = \left| \sum_{i=0}^n p_i z^i \right| \leq \sum_{i=0}^n |p_i| |z|^i = \sum_{i=0}^n |p_i| |z|^i \leq \sum_{i=0}^n |p_i| |z|^n = M|z|^n.$$

Analog haben wir

$$|p_n z^n - p(z)| = \left| - \sum_{i=0}^{n-1} p_i z^i \right| = \left| \sum_{i=0}^{n-1} p_i z^i \right| \leq \sum_{i=0}^{n-1} |p_i| |z|^{n-1} \leq M|z|^{n-1} \leq \frac{|p_n|}{2} |z|^n,$$

wobei wir am Ende $|z| \geq R \geq 2M/|p_n|$, d.h. $M \leq \frac{|p_n|}{2}|z|$, verwendet haben. Zusammen mit der Dreiecksungleichung,

$$|p_n| |z|^n = |p_n z^n| = |p_n z^n - p(z) + p(z)| \leq |p_n z^n - p(z)| + |p(z)|,$$

folgt daraus

$$|p(z)| \geq |p_n| |z|^n - |p_n z^n - p(z)| \geq |p_n| |z|^n - \frac{|p_n|}{2} |z|^n = \frac{|p_n|}{2} |z|^n = m|z|^n.$$

Damit ist also die Abschätzung (VI.11) gezeigt.

Daraus werden wir nun ableiten, dass $z_0 \in \mathbb{C}$ existiert, sodass

$$|p(z)| \geq |p(z_0)|, \quad \text{für alle } z \in \mathbb{C}. \quad (\text{VI.12})$$

Dies bedeutet gerade, dass die stetige Funktion $f: \mathbb{C} \rightarrow \mathbb{R}$, $f(z) := |p(z)|$, ein globales Minimum besitzt. Sei dazu $\tilde{R} := \max\{R, \sqrt[n]{|p(0)|/m}\}$. Da die abgeschlossene Scheibe $\{z \in \mathbb{C} : |z| \leq \tilde{R}\}$ kompakt ist, nimmt f darauf ein Minimum an, d.h. es existiert $z_0 \in \mathbb{C}$, sodass

$$|p(z)| \geq |p(z_0)|, \quad \text{für alle } z \in \mathbb{C} \text{ mit } |z| \leq \tilde{R}.$$

Es gilt aber auch

$$|p(z)| \geq |p(z_0)|, \quad \text{für alle } z \in \mathbb{C} \text{ mit } |z| \geq \tilde{R},$$

denn für diese z folgt aus (VI.11): $|p(z)| \geq m|z|^n \geq m\tilde{R}^n \geq |p(0)| \geq |p(z_0)|$. Damit ist also (VI.12) bewiesen.

Wir werden nun $p(z_0) = 0$ zeigen, und damit den Beweis des Satzes abschließen. Wir gehen indirekt vor und nehmen $p(z_0) \neq 0$ an. Beachte, dass dann auch $\tilde{p}(z) = p(z + z_0)/p(z_0)$ ein nicht-konstantes Polynom vom Grad n darstellt, es ist daher von der Form

$$\tilde{p}(z) = 1 + \tilde{p}_k z^k + \cdots + \tilde{p}_n z^n,$$

wobei $k \geq 1$, $\tilde{p}_i \in \mathbb{C}$ und $\tilde{p}_k \neq 0$. Nach (VI.12) gilt für jedes $z \in \mathbb{C}$:

$$|\tilde{p}(z)| = \frac{|p(z + z_0)|}{|p(z_0)|} \geq \frac{|p(z_0)|}{|p(z_0)|} = 1.$$

Sei nun $w \in \mathbb{C}$ mit $w^k = -1/\tilde{p}_k$,⁵ und betrachte das Polynom $q(z) := \tilde{p}(wz)$. Nach Konstruktion ist q von der Form

$$q(z) = 1 - z^k + q_{k+1}z^{k+1} + \cdots + q_n z^n, \quad (\text{VI.13})$$

wobei $q_i \in \mathbb{C}$. Aus der entsprechenden Eigenschaft von \tilde{p} folgt sofort

$$|q(z)| \geq 1, \quad \text{für alle } z \in \mathbb{C}. \quad (\text{VI.14})$$

Sei nun $C := \sum_{i=k+1}^n |q_i|$ und $\rho := \min\{1, 1/2C\}$. Für jedes $z \in \mathbb{C}$ mit $|z| \leq \rho$ gilt daher $|z| \leq 1$ und $C|z| \leq 1/2$, folglich

$$\left| q_{k+1}z^{k+1} + \cdots + q_n z^n \right| \leq \sum_{i=k+1}^n |q_i||z|^i \leq \sum_{i=k+1}^n |q_i||z|^{k+1} = C|z|^{k+1} \leq \frac{1}{2}|z|^k.$$

Zusammen mit (VI.13) erhalten wir aus der Dreiecksungleichung,

$$|q(z)| \leq |1 - z^k| + \frac{1}{2}|z|^k, \quad \text{für alle } z \in \mathbb{C} \text{ mit } |z| \leq \rho.$$

Für reelle $t \in (0, \rho]$ folgt daraus

$$|q(t)| \leq |1 - t^k| + \frac{1}{2}t^k = 1 - t^k + \frac{1}{2}t^k = 1 - \frac{1}{2}t^k < 1,$$

im Widerspruch zu (VI.14).

Da die Annahme $p(z_0) \neq 0$ auf einen Widerspruch führt, muss also $p(z_0) = 0$ gelten. Somit hat p eine Nullstelle und der Beweis ist vollständig. \square

⁵Jedes $z \in \mathbb{C}$ besitzt eine k -te Wurzel, d.h. es existiert $w \in \mathbb{C}$ mit $w^k = z$. Schreiben wir $z = r e^{i\theta}$, wobei $r \geq 0$ und $\theta \in \mathbb{R}$, dann hat $w = \sqrt[k]{r} e^{i\theta/k}$ die gewünschte Eigenschaft.

VI.2.11. BEMERKUNG (Algebraischer Abschluss). Zu jedem Körper \mathbb{K} existiert ein algebraisch abgeschlossener Körper $\hat{\mathbb{K}}$, der \mathbb{K} enthält und für den die Inklusion $\mathbb{K} \rightarrow \hat{\mathbb{K}}$ mit Addition und Multiplikation verträglich ist. Dieser, i.W. eindeutige, Körper $\hat{\mathbb{K}}$ wird der *algebraische Abschluss* von \mathbb{K} genannt. Aus dem Fundamentalsatz der Algebra folgt, dass der Körper \mathbb{C} mit dem algebraischen Abschluss des Körpers \mathbb{R} übereinstimmt, in Zeichen: $\mathbb{C} = \hat{\mathbb{R}}$. Der algebraische Abschluss von \mathbb{Q} ist der Körper der algebraischen Zahlen.

Wir kommen nun zur Definition des charakteristischen Polynoms einer quadratischen Matrix. Wir wollen jeder Matrix $A \in M_{n \times n}(\mathbb{K})$ ein Polynom $p_A \in \mathbb{K}[z]$ zuordnen, sodass

$$p_A(\lambda) = \det(A - \lambda I_n),$$

für jedes $\lambda \in \mathbb{K}$. Aus der Leibniz'schen Formel folgt sofort, dass $\lambda \mapsto \det(A - \lambda I_n)$ eine *Polynomfunktion* bildet, d.h. es existieren $p_i \in \mathbb{K}$ mit

$$\det(A - \lambda I_n) = p_0 + p_1 \lambda + \cdots + p_n \lambda^n,$$

für jedes $\lambda \in \mathbb{K}$. Die Koeffizienten p_i sind durch diese Bedingung aber i.A. nicht eindeutig bestimmt, das Polynom p_A kann daher nicht einfach durch $p_0 + p_1 z + \cdots + p_n z^n$ definiert werden! Betrachten wir etwa den Körper \mathbb{Z}_2 und das Polynom $p = z + z^2 \in \mathbb{Z}_2[z]$, dann gilt $p(\lambda) = 0$, für jedes $\lambda \in \mathbb{Z}_2$, obwohl $0 \neq p \in \mathbb{Z}_2[z]$. Für unendliche Körper ist die Abbildung $\mathbb{K}[z] \rightarrow F(\mathbb{K}, \mathbb{K})$, die einem Polynom p die Polynomfunktion $\lambda \mapsto p(\lambda)$ zuordnet, injektiv, siehe Korollar IV.6.27. Für unendliche Körper kann das charakteristische Polynom also durch $p_A(\lambda) = \det(A - \lambda I_n)$, $\lambda \in \mathbb{K}$, definiert werden, es ist dadurch eindeutig bestimmt. I.A. müssen wir vorsichtiger vorgehen.

Ist $P \in M_{n \times n}(\mathbb{K}[z])$ eine Matrix von Polynomen, dann definieren wir deren Determinante, $\det(P) \in \mathbb{K}[z]$, durch die Leibniz-Formel,

$$\det(P) := \sum_{\sigma \in \mathfrak{S}_n} \operatorname{sgn}(\sigma) P_{1,\sigma(1)} \cdots P_{n,\sigma(n)} \in \mathbb{K}[z],$$

wobei $P_{ij} \in \mathbb{K}[z]$ das Polynom in der i -ten Zeile und j -ten Spalte von P bezeichnet. Aus Satz V.3.6 folgt

$$(\det(P))(\lambda) = \det(P(\lambda)), \quad \lambda \in \mathbb{K}, \quad (\text{VI.15})$$

wobei $P(\lambda) \in M_{n \times n}(\mathbb{K})$ die Matrix mit Eintragungen $P(\lambda)_{ij} = P_{ij}(\lambda)$ bezeichnet. Für beliebige Matrizen von Polynomen, $P, Q \in M_{n \times n}(\mathbb{K}[z])$, gilt

$$\det(PQ) = \det(P) \det(Q) \in \mathbb{K}[z], \quad (\text{VI.16})$$

wobei das Produkt von Polynommatrizen analog zum gewöhnlichen Matrizenprodukt definiert ist, $(PQ)_{ij} = \sum_k P_{ik} Q_{kj} \in \mathbb{K}[z]$. Über unendlichen Körpern folgt dies aus (VI.15), für beliebige Körper siehe Aufgabe 49. Beachte, dass auch $M_{n \times n}(\mathbb{K}[z])$ eine assoziative Algebra bildet. Indem wir Matrizen mit Eintragungen in \mathbb{K} als Matrizen konstanter Polynome auffassen erhalten wir $M_{n \times n}(\mathbb{K}) \subseteq$

$M_{n \times n}(\mathbb{K}[z])$, und dies ist mit Addition, Skalarmultiplikation, Matrizenmultiplikation und der Determinante verträglich.

VI.2.12. DEFINITION (Charakteristisches Polynom einer Matrix). Unter dem *charakteristischen Polynom* einer Matrix $A \in M_{n \times n}(\mathbb{K})$ verstehen wir das Polynom $p_A \in \mathbb{K}[z]$,

$$p_A := \det(A - zI_n) = \sum_{\sigma \in \mathfrak{S}_n} \text{sgn}(\sigma)(A - zI_n)_{1,\sigma(1)} \cdots (A - zI_n)_{n,\sigma(n)}. \quad (\text{VI.17})$$

Dabei bezeichnet $(A - zI_n)_{ij} \in \mathbb{K}[z]$ das Polynom in der i -ten Zeile und j -ten Spalte von $A - zI_n \in M_{n \times n}(\mathbb{K}[z])$. Für $i \neq j$ ist $(A - zI_n)_{ij} = A_{ij}$ ein konstantes Polynom und $(A - zI_n)_{ii} = A_{ii} - z$ hat Grad 1.

VI.2.13. BEISPIEL. Betrachte die Matrix $A = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \in M_{2 \times 2}(\mathbb{Z}_2)$ über dem Körper $\mathbb{K} = \mathbb{Z}_2$. Für ihr charakteristische Polynom $p_A \in \mathbb{Z}_2[z]$ gilt

$$p_A = \det(A - zI_2) = \det \begin{pmatrix} 0 - z & 0 \\ 0 & 1 - z \end{pmatrix} = -z(1 - z) = z + z^2.$$

Beachte $p_A(\lambda) = 0$, für jedes $\lambda \in \mathbb{Z}_2$, denn $p_A(0) = 0 + 0^2 = 0$ und $p_A(1) = 1 + 1^2 = 0$. Das charakteristische Polynom ist jedoch nicht trivial, $0 \neq p_A \in \mathbb{Z}_2[z]$.

VI.2.14. LEMMA. Sei $A \in M_{n \times n}(\mathbb{K})$. Dann gilt:

- (a) $p_A(\lambda) = \det(A - \lambda I_n)$, für alle $\lambda \in \mathbb{K}$.
- (b) $\deg(p_A) = n$. Bezeichnen wir die Koeffizienten mit $p_A = p_0 + p_1 z + \cdots + p_n z^n$, dann gilt $p_n = (-1)^n$, $p_{n-1} = (-1)^{n-1} \text{tr}(A)$ und $p_0(A) = \det(A)$.
- (c) $p_{S^{-1}AS} = p_A$, für jedes $S \in \text{GL}_n(\mathbb{K})$.
- (d) Ist $A \in M_{n \times n}(\mathbb{K})$, $B \in M_{m \times m}(\mathbb{K})$ und $C \in M_{(n+m) \times (n+m)}(\mathbb{K})$ eine Blockmatrix der Form $C = \begin{pmatrix} A & * \\ 0 & B \end{pmatrix}$, dann gilt $p_C = p_A p_B$.
- (e) Für das charakteristische Polynom einer Dreiecksmatrix,

$$D = \begin{pmatrix} \lambda_1 & * & \cdots & * \\ & \lambda_2 & \ddots & \vdots \\ & & \ddots & * \\ & & & \lambda_n \end{pmatrix},$$

gilt $p_D = (\lambda_1 - z)(\lambda_2 - z) \cdots (\lambda_n - z)$.

BEWEIS. Behauptung (a) folgt aus (VI.15), denn

$$p_A(\lambda) = (\det(A - zI_n))(\lambda) = \det(A - \lambda I_n).$$

Daraus erhalten wir auch $p_0 = p_A(0) = \det(A - 0I_n) = \det(A)$. Weiters hat

$$(A - zI_n)_{1,\sigma(1)} \cdots (A - zI_n)_{n,\sigma(n)}$$

höchstens Grad $n - 2$, falls $\text{id} \neq \sigma \in \mathfrak{S}_n$. Für $\text{id} = \sigma \in \mathfrak{S}_n$ gilt

$$(A - zI_n)_{1,\sigma(1)} \cdots (A - zI_n)_{n,\sigma(n)} = (-1)^n z^n + (-1)^{n-1} (A_{11} + \cdots + A_{nn}) z^{n-1} + q,$$

wobei $q \in \mathbb{K}[z]$ und $\deg(q) \leq n - 2$. Mit (VI.17) erhalten wir also $\deg(p_A) = n$, $p_n = (-1)^n$ und $p_{n-1} = (-1)^{n-1} \operatorname{tr}(A)$. Behauptung (c) folgt aus (VI.16), denn

$$\begin{aligned} p_{S^{-1}AS} &= \det(S^{-1}AS - zI_n) = \det(S^{-1}(A - zI_n)S) \\ &= \det(S)^{-1} \det(A - zI_n) \det(S) = \det(A - zI_n) = p_A. \end{aligned}$$

Um (d) einzusehen, beobachten wir zunächst, dass

$$(C - zI)_{1,\sigma(1)} \cdots (C - zI)_{n+m,\sigma(n+m)} = 0,$$

falls ein $i > n$ existiert, für das $\sigma(i) \leq n$ gilt. In der Definition von p_C ist daher nur über jene Permutationen $\sigma \in \mathfrak{S}_{n+m}$ zu summieren, für die $\sigma(\{n+1, \dots, n+m\}) \subseteq \{n+1, \dots, n+m\}$ gilt. Solche Permutationen müssen aber auch $\sigma(\{1, \dots, n\}) \subseteq \{1, \dots, n\}$ genügen und können daher mit $\mathfrak{S}_n \times \mathfrak{S}_m$ identifiziert werden. Somit

$$\begin{aligned} p_C &= \sum_{\sigma \in \mathfrak{S}_{n+m}} \operatorname{sgn}(\sigma) (C - zI)_{1,\sigma(1)} \cdots (C - zI)_{n+m,\sigma(n+m)} \\ &= \sum_{\tau \in \mathfrak{S}_n, \pi \in \mathfrak{S}_m} \operatorname{sgn}(\tau) \operatorname{sgn}(\pi) \prod_{i=1}^n (C - zI)_{i,\tau(i)} \prod_{j=1}^m (C - zI)_{n+j,n+\pi(j)} \\ &= \sum_{\tau \in \mathfrak{S}_n, \pi \in \mathfrak{S}_m} \operatorname{sgn}(\tau) \operatorname{sgn}(\pi) \prod_{i=1}^n (A - zI)_{i,\tau(i)} \prod_{j=1}^m (B - zI)_{j,\pi(j)} \\ &= \sum_{\tau \in \mathfrak{S}_n} \operatorname{sgn}(\tau) \prod_{i=1}^n (A - zI)_{i,\tau(i)} \sum_{\pi \in \mathfrak{S}_m} \operatorname{sgn}(\pi) \prod_{j=1}^m (B - zI)_{j,\pi(j)} = p_A p_B \end{aligned}$$

Behauptung (e) lässt sich mittels Induktion nach n aus (d) herleiten. Wir geben einen anderen, direkteren Beweis. Beachte, dass wegen der Dreiecksform von D offensichtlich $(D - zI_n)_{1,\sigma(1)} \cdots (D - zI_n)_{n,\sigma(n)} = 0$, für jedes $\operatorname{id} \neq \sigma \in \mathfrak{S}_n$. Aus (VI.17) folgt daher

$$p_D = (D - zI_n)_{11} \cdots (D - zI_n)_{nn} = (\lambda_1 - z) \cdots (\lambda_n - z),$$

womit auch (e) gezeigt wäre. \square

VI.2.15. BEMERKUNG. Nach Lemma VI.2.14(c) haben ähnliche Matrizen das selbe charakteristische Polynom, d.h. jeder Koeffizient p_i des charakteristischen Polynoms $p_A = p_0 + p_1 z + \cdots + p_n z^n$ einer Matrix $A \in M_{n \times n}(\mathbb{K})$ bleibt unverändert, wenn wir A durch eine ähnliche Matrix $B = S^{-1}AS$, $S \in \operatorname{GL}_n(\mathbb{K})$ ersetzen. Für die Koeffizienten p_{n-1} und p_0 bedeutet dies gerade $\operatorname{tr}(S^{-1}AS) = \operatorname{tr}(A)$ bzw. $\det(S^{-1}AS) = \det(A)$, vgl. Lemma VI.2.14(b).

VI.2.16. DEFINITION (Charakteristisches Polynom eines Endomorphismus). Sei V ein endlich dimensionaler \mathbb{K} -Vektorraum und $\varphi: V \rightarrow V$ linear. Unter dem *charakteristischen Polynom von φ* verstehen wir das charakteristische Polynom der Matrix $[\varphi]_{BB}$, wobei B eine geordnete Basis von V bezeichnet. Beachte, dass

dies nach Lemma VI.2.14(c) nicht von der Basis B abhängt und daher wohldefiniert ist. Wir bezeichnen das charakteristische Polynom von φ mit $p_\varphi \in \mathbb{K}[z]$.

VI.2.17. BEMERKUNG. Ist $A \in M_{n \times n}(\mathbb{K})$ und bezeichnet $\varphi: \mathbb{K}^n \rightarrow \mathbb{K}^n$, $\varphi(x) = Ax$, die damit assoziierte lineare Abbildung, dann gilt $p_\varphi = p_A$, denn bezüglich der Standardbasis B von \mathbb{K}^n gilt $[\varphi]_{BB} = A$.

VI.2.18. PROPOSITION. Sei V ein n -dimensionaler \mathbb{K} -Vektorraum, $\varphi: V \rightarrow V$ linear und $p_\varphi \in \mathbb{K}[z]$ das charakteristische Polynom von φ . Dann gilt:

- (a) $p_\varphi(\lambda) = \det(\varphi - \lambda \text{id}_V)$, für alle $\lambda \in \mathbb{K}$.
- (b) Die Eigenwerte von φ sind genau die Nullstellen von p_φ , für das Spektrum von φ gilt daher: $\sigma(\varphi) = \{\lambda \in \mathbb{K} : p_\varphi(\lambda) = 0\}$.
- (c) $\deg(p_\varphi) = n$. Bezeichnen wir die Koeffizienten mit $p_\varphi = p_0 + p_1 z + \dots + p_n z^n$, dann gilt $p_n = (-1)^n$, $p_{n-1} = (-1)^{n-1} \text{tr}(\varphi)$ und $p_0 = \det(\varphi)$.
- (d) $p_{\psi^{-1}\varphi\psi} = p_\varphi$, für jeden linearen Isomorphismus $\psi: W \rightarrow V$.
- (e) $p_{\lambda \text{id}_V} = (\lambda - z)^n$, für jedes $\lambda \in \mathbb{K}$.
- (f) Sei W ein Teilraum von V , sodass $\varphi(W) \subseteq W$. Bezeichnet $\varphi|_W: W \rightarrow W$ die Einschränkung und $\bar{\varphi}: V/W \rightarrow V/W$, $\bar{\varphi}([v]) = [\varphi(v)]$ die induzierte lineare Abbildung auf dem Quotientenraum, dann gilt $p_\varphi = p_{\varphi|_W} p_{\bar{\varphi}}$.

BEWEIS. Behauptung (a) folgt sofort aus Lemma VI.2.14(a). Ist nämlich B eine Basis von V und $\lambda \in \mathbb{K}$, dann gilt

$$p_\varphi(\lambda) = p_{[\varphi]_{BB}}(\lambda) = \det([\varphi]_{BB} - \lambda I_n) = \det([\varphi - \lambda \text{id}_V]_{BB}) = \det(\varphi - \lambda \text{id}_V).$$

Zusammen mit Proposition VI.1.3 erhalten wir daraus auch (b). Aussagen (c) folgt aus Lemma VI.2.14(b), denn $\det([\varphi]_{BB}) = \det(\varphi)$ und $\text{tr}([\varphi]_{BB}) = \text{tr}(\varphi)$, vgl. Aufgabe 23. Behauptung (d) folgt aus Lemma VI.2.14(c), denn für jede Basis C von W ist $S := [\psi]_{BC} \in \text{GL}_n(\mathbb{K})$ und $[\psi^{-1}\varphi\psi]_{CC} = [\psi]_{CB}[\varphi]_{BB}[\psi]_{BC} = S^{-1}[\varphi]_{BB}S$, also

$$p_{\psi^{-1}\varphi\psi} = p_{[\psi^{-1}\varphi\psi]_{CC}} = p_{S^{-1}[\varphi]_{BB}S} = p_{[\varphi]_{BB}} = p_\varphi.$$

Behauptung (e) folgt aus Lemma VI.2.14(e), denn $[\lambda \text{id}_V]_{BB} = \lambda I_n$, also $p_{\lambda \text{id}_V} = p_{[\lambda \text{id}_V]_{BB}} = p_{\lambda I_n} = (\lambda - z)^n$. Es bleibt daher nur noch (f) zu zeigen. Wir wählen dazu eine Basis $B' = (b_1, \dots, b_m)$ von W und ergänzen sie zu einer Basis $B = (b_1, \dots, b_m, b_{m+1}, \dots, b_n)$ von V . Bezeichnet $\pi: V \rightarrow V/W$ die kanonische Projektion, dann bildet $\bar{B} := (\pi(b_{m+1}), \dots, \pi(b_n))$ eine Basis von V/W , siehe Korollar II.6.7, und es gilt:

$$[\varphi]_{BB} = \begin{pmatrix} [\varphi|_W]_{B'B'} & * \\ 0 & [\bar{\varphi}]_{\bar{B}\bar{B}} \end{pmatrix}. \quad (\text{VI.18})$$

Um dies einzusehen beginnen wir mit der Definition der Matrix $[\varphi]_{BB}$,

$$\varphi(b_j) = \sum_{i=1}^n ([\varphi]_{BB})_{ij} b_i. \quad (\text{VI.19})$$

Für $j \leq m$ ist $b_j \in W$ und nach Voraussetzung an φ daher auch $\varphi(b_j) \in W$. Daraus folgt

$$([\varphi]_{BB})_{ij} = \begin{cases} ([\varphi|_W]_{B'B'})_{ij} & \text{für } j \leq m \text{ und } i \leq m \\ 0 & \text{für } j \leq m \text{ und } i > m. \end{cases}$$

Dies erklärt die linken beiden Blöcke in (VI.18). Sei nun $j > m$. Wenden wir auf (VI.19) die Projektion $\pi: V \rightarrow V/W$ an, so folgt

$$\bar{\varphi}(\pi(b_j)) = \pi(\varphi(b_j)) = \sum_{i=1}^n ([\varphi]_{BB})_{ij} \pi(b_i) = \sum_{i=m+1}^n ([\varphi]_{BB})_{ij} \pi(b_i),$$

denn $\pi(b_i) = 0$, für alle $i \leq m$. Daraus erhalten wir $([\varphi]_{BB})_{ij} = ([\bar{\varphi}]_{\bar{B}\bar{B}})_{i-m, j-m}$ für alle $i, j > m$, also den rechten unteren Block in (VI.18). Mit Lemma VI.2.14(d) erhalten wir aus (VI.18) nun $p_\varphi = p_{[\varphi]_{BB}} = p_{[\varphi|_W]_{B'B'}} p_{[\bar{\varphi}]_{\bar{B}\bar{B}}} = p_{\varphi|_W} p_{\bar{\varphi}}$. \square

VI.2.19. DEFINITION (Algebraische Vielfachheit von Eigenwerten). Sei V ein endlich-dimensionaler \mathbb{K} -Vektorraum und $\varphi: V \rightarrow V$ linear. Unter der *algebraischen Vielfachheit* eines Eigenwerts λ von φ verstehen wir die Vielfachheit der Nullstelle λ des charakteristischen Polynoms p_φ , vgl. Proposition VI.2.18(b). Für Matrizen $A \in M_{n \times n}(\mathbb{K})$ definieren wir die algebraische Vielfachheit der Eigenwerte über die assoziierte lineare Abbildung, $\mathbb{K}^n \rightarrow \mathbb{K}^n$, $x \mapsto Ax$. Offensichtlich stimmt dies mit der algebraischen Vielfachheit der Nullstelle λ von p_A überein.

VI.2.20. BEISPIEL. Algebraische und geometrische Vielfachheit eines Eigenwerts sind i.A. verschieden. Betrachte dazu die Matrix

$$A = \begin{pmatrix} 3 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 \\ 0 & 0 & 5 & 1 \\ 0 & 0 & 0 & 5 \end{pmatrix}$$

mit charakteristischem Polynom $p_A = (3 - z)^2(5 - z)^2$. Diese Matrix hat zwei Eigenwerte, $\lambda_1 = 3$ und $\lambda_2 = 5$. Beide haben algebraische Vielfachheit 2. Auch die geometrische Vielfachheit des Eigenwerts $\lambda_1 = 3$ ist 2, denn der zugehörige Eigenraum wird von den beiden Einheitsvektoren e_1 und e_2 aufgespannt, $E_{\lambda_1} = \ker(A - 3I_4) = \langle e_1, e_2 \rangle$. Die geometrische Vielfachheit des Eigenwerts $\lambda_2 = 5$ ist allerdings nur 1, denn der entsprechende Eigenraum $\ker(A - 5I_4) = \langle e_3 \rangle$ ist 1-dimensional.

VI.2.21. PROPOSITION. Sei V ein endlich-dimensionaler \mathbb{K} -Vektorraum und $\varphi: V \rightarrow V$ linear. Dann ist die algebraische Vielfachheit jedes Eigenwerts von φ mindestens so groß wie seine geometrische Vielfachheit.

BEWEIS. Sei λ ein Eigenwert von φ , bezeichne $W = \ker(\varphi - \lambda \text{id}_V)$ den entsprechenden Eigenraum und $m = \dim(W)$ die geometrische Vielfachheit von λ .

Es gilt daher $\varphi|_W = \lambda \text{id}_W$, insbesondere ist W invariant unter φ . Es bezeichne $\bar{\varphi}: V/W \rightarrow V/W$, $\bar{\varphi}([v]) = [\varphi(v)]$, die induzierte lineare Abbildung. Nach Proposition VI.2.18(e)&(f) gilt für die charakteristischen Polynome

$$p_\varphi = p_{\varphi|_W} p_{\bar{\varphi}} = (\lambda - z)^m p_{\bar{\varphi}},$$

die algebraische Vielfachheit von λ ist daher mindestens m . \square

Aus der vorangehenden Proposition folgt sofort, dass auch für quadratische Matrizen die algebraische Vielfachheit jedes Eigenwerts mindestens so groß wie seine geometrische Vielfachheit ist.

VI.2.22. SATZ. *Sei V ein endlich-dimensionaler \mathbb{K} -Vektorraum. Eine lineare Abbildung $\varphi: V \rightarrow V$ ist genau dann diagonalisierbar, wenn ihr charakteristisches Polynom über \mathbb{K} in Linearfaktoren zerfällt und die algebraische Vielfachheit jedes Eigenwerts mit seiner geometrischen Vielfachheit übereinstimmt.*

BEWEIS. Es bezeichne p das charakteristische Polynom von φ . Zerfällt p in Linearfaktoren, dann stimmt die Summe der algebraischen Vielfachheiten aller Nullstellen von p mit $\deg(p) = \dim(V)$ überein. Somit ist die Summe der algebraischen Vielfachheiten aller Eigenwerte von φ gleich $\dim(V)$. Stimmen darüber hinaus die algebraischen mit den geometrischen Vielfachheit überein, dann ist also auch die Summe der geometrischen Vielfachheiten aller Eigenwerte gleich $\dim(V)$. Nach Satz VI.1.18 ist φ in diesem Fall diagonalisierbar. Die umgekehrte Implikation folgt sofort aus Lemma VI.2.14(e). \square

VI.2.23. DEFINITION (Triangulierbarkeit). Sei V ein endlich-dimensionaler \mathbb{K} -Vektorraum. Eine lineare Abbildung $\varphi: V \rightarrow V$ wird *triangulierbar* genannt, falls eine geordnete Basis B von V existiert, sodass die Matrixdarstellung von φ bezüglich B obere Dreiecksgestalt hat, d.h.

$$[\varphi]_{BB} = \begin{pmatrix} \lambda_1 & * & \cdots & * \\ 0 & \lambda_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & * \\ 0 & \cdots & 0 & \lambda_n \end{pmatrix}.$$

Eine quadratische Matrix $A \in M_{n \times n}(\mathbb{K})$ wird triangulierbar genannt, wenn die damit assoziierte lineare Abbildung $\mathbb{K}^n \rightarrow \mathbb{K}^n$, $x \mapsto Ax$, triangulierbar ist.

Eine quadratische Matrix ist genau dann triangulierbar, wenn sie ähnlich zu einer oberen Dreiecksmatrix ist. Eine lineare Abbildung $\varphi: V \rightarrow V$ ist genau dann triangulierbar wenn ihre Matrixdarstellung bezüglich einer (und dann jeder) Basis triangulierbar ist. Offensichtlich ist jede diagonalisierbare lineare Abbildung auch triangulierbar. Umgekehrt muss eine triangulierbare Abbildung bzw. Matrix nicht diagonalisierbar sein. Beispielsweise ist die Matrix $\begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix}$ triangulierbar aber nicht diagonalisierbar, vgl. Beispiel VI.1.12. Nicht jede Matrix ist triangulierbar. Etwa hat die reelle Matrix $\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ keinen einzigen reellen Eigenwert und kann daher über \mathbb{R} nicht triangulierbar sein.

VI.2.24. SATZ. Sei V ein endlich-dimensionaler \mathbb{K} -Vektorraum. Eine lineare Abbildung $\varphi: V \rightarrow V$ ist genau dann triangulierbar, wenn ihr charakteristisches Polynom über \mathbb{K} in Linearfaktoren zerfällt.

BEWEIS. Aus Lemma VI.2.14(e) folgt sofort, dass das charakteristische Polynom einer triangulierbaren Matrix in Linearfaktoren zerfällt. Für die umgekehrte Implikation sei nun φ eine lineare Abbildung, deren charakteristisches Polynom in Linearfaktoren zerfällt. Wir werden nun mittels Induktion nach der Dimension von V zeigen, dass φ triangulierbar ist. Der Induktionsanfang, $\dim(V) = 1$, ist trivial, denn jede 1×1 -Matrix hat obere Dreiecksgestalt. Für den Induktionsschritt sei nun $\dim(V) \geq 2$. Da das charakteristische Polynom p_φ in Linearfaktoren zerfällt, existiert eine Nullstelle $\lambda \in \mathbb{K}$, d.h. $p_\varphi(\lambda) = 0$. Nach Proposition VI.2.18(b) ist λ Eigenwert von φ . Sei $0 \neq b_1 \in V$ ein entsprechender Eigenvektor, d.h. $\varphi(b_1) = \lambda b_1$, und bezeichne $W := \langle b_1 \rangle$ den davon aufgespannten 1-dimensionalen Teilraum. Nach Proposition VI.2.18(f) gilt

$$p_\varphi = p_{\varphi|_W} p_{\bar{\varphi}} = (\lambda - z) p_{\bar{\varphi}},$$

wobei $p_{\bar{\varphi}}$ das charakteristische Polynom der linearen Abbildung $\bar{\varphi}: V/W \rightarrow V/W$, $\bar{\varphi}([v]) = [\varphi(v)]$, bezeichnet. Daraus schließen wir, dass auch $p_{\bar{\varphi}}$ in Linearfaktoren zerfällt. Nach Induktionsvoraussetzung ist $\bar{\varphi}$ also triangulierbar, denn $\dim(V/W) = \dim(V) - 1 < \dim(V)$. Es existiert daher eine geordnete Basis $\bar{B} = (\bar{b}_2, \dots, \bar{b}_n)$ von V/W , sodass $[\bar{\varphi}]_{\bar{B}\bar{B}}$ obere Dreiecksgestalt hat:

$$[\bar{\varphi}]_{\bar{B}\bar{B}} = \begin{pmatrix} * & \cdots & * \\ & \ddots & \vdots \\ & & * \end{pmatrix}.$$

Wählen wir $b_2, \dots, b_n \in V$ mit $[b_i] = \bar{b}_i$, dann bildet $B = (b_1, b_2, \dots, b_n)$ eine Basis von V , siehe Korollar II.6.7. Nach Konstruktion hat $[\varphi]_{BB}$ die Form

$$[\varphi]_{BB} = \begin{pmatrix} \lambda & * \\ 0 & [\bar{\varphi}]_{\bar{B}\bar{B}} \end{pmatrix} = \begin{pmatrix} \lambda & * & \cdots & * \\ & * & \cdots & * \\ & & \ddots & \vdots \\ & & & * \end{pmatrix},$$

denn $\varphi(b_1) = \lambda b_1 + 0b_2 + \cdots + 0b_n$ und, für $j \geq 2$,

$$\varphi(\bar{b}_j) = \bar{\varphi}([b_j]) = [\varphi(b_j)] = \left[\sum_{i=1}^n ([\varphi]_{BB})_{ij} b_i \right] = \sum_{i=1}^n ([\varphi]_{BB})_{ij} [b_i] = \sum_{i=2}^n ([\varphi]_{BB})_{ij} \bar{b}_i,$$

da $[b_1] = 0 \in V/W$, folglich $([\varphi]_{BB})_{ij} = ([\bar{\varphi}]_{\bar{B}\bar{B}})_{i-1, j-1}$, für alle $i, j \geq 2$. Somit ist φ also triangulierbar. \square

VI.2.25. KOROLLAR. Sei V ein n -dimensionaler Vektorraum über \mathbb{K} und $\varphi: V \rightarrow V$ eine lineare Abbildung deren charakteristisches Polynom in Linearfaktoren zerfällt, d.h.

$$p_\varphi = p_0 + p_1 z + \cdots + p_n z^n = (\lambda_1 - z) \cdots (\lambda_n - z), \quad (\text{VI.20})$$

wobei $\lambda_1, \dots, \lambda_n$ die Eigenwerte von φ bezeichnen und ihrer algebraischen Vielfachheit entsprechend oft aufgelistet sind. Dann gilt

$$\sigma(q(\varphi)) = q(\sigma(\varphi))$$

für jedes Polynom $q \in \mathbb{K}[z]$. Weiters ist

$$(-1)^{n-j} p_{n-j} = \sum_{1 \leq i_1 < \cdots < i_j \leq n} \lambda_{i_1} \cdots \lambda_{i_j},$$

für jedes $1 \leq j \leq n$. Insbesondere gilt

$$\text{tr}(\varphi) = \lambda_1 + \cdots + \lambda_n \quad \text{und} \quad \det(\varphi) = \lambda_1 \cdots \lambda_n,$$

d.h. die Summe der Eigenwerte stimmt mit der Spur überein und ihr Produkt liefert die Determinante.

BEWEIS. Nach Satz VI.2.24 ist φ triangulierbar. Die erste Behauptung über das Spektrum von $q(\varphi)$ folgt daher aus Proposition VI.1.23. Die Formeln für p_j erhalten wir durch Ausmultiplizieren der rechten Seite von (VI.20). Für $j = n$ erhalten wir daraus $\det(\varphi) = p_0 = \lambda_1 \cdots \lambda_n$, und für $j = 1$ auch $\text{tr}(\varphi) = (-1)^{n-1} p_{n-1} = \lambda_1 + \cdots + \lambda_n$, vgl. Proposition VI.2.18(c). Diese Formeln für die Spur und die Determinante lassen sich, nach Wahl einer Basis in der φ obere Dreiecksgestalt hat, auch direkt ablesen, denn bezüglich einer solchen Basis stimmen die Diagonaleinträge mit den Eigenwerten überein. \square

Für algebraisch abgeschlossene Körper erhalten wir daher:

VI.2.26. KOROLLAR. Sei V ein n -dimensionaler Vektorraum über einem algebraisch abgeschlossenen Körper \mathbb{K} . Dann ist jede lineare Abbildung $\varphi: V \rightarrow V$ triangulierbar. Weiters gilt $\sigma(q(\varphi)) = q(\sigma(\varphi))$, für jedes Polynom $q \in \mathbb{K}[z]$. Bezeichnen $\lambda_1, \dots, \lambda_n$ die Eigenwerte von φ , ihrer algebraischen Vielfachheit entsprechend oft angeführt, dann gilt $\text{tr}(\varphi) = \lambda_1 + \cdots + \lambda_n$ sowie $\det(\varphi) = \lambda_1 \cdots \lambda_n$.

VI.3. Jordan'sche Normalform. Sei V ein endlich dimensionaler Vektorraum über \mathbb{K} , und $\varphi: V \rightarrow V$ eine lineare Abbildung, deren charakteristisches Polynom in Linearfaktoren zerfällt. Nach Satz VI.2.24 ist φ daher triangulierbar, d.h. es existiert eine Basis B von V , sodass $[\varphi]_{BB}$ obere Dreiecksgestalt hat. In diesem Abschnitt werden wir sehen, dass sich die Matrix $[\varphi]_{BB}$ durch geschickte Wahl der Basis B noch erheblich vereinfachen lässt. Wie wir bereits weiter oben am Beispiel $A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ beobachtet haben, wird φ i.A. nicht diagonalisierbar sein, d.h. es ist nicht immer möglich $[\varphi]_{BB}$ auf Diagonalgestalt zu bringen. Wir beginnen mit der sogenannten Primärzerlegung.

VI.3.1. DEFINITION (Verallgemeinerte Eigenräume). Sei V ein endlich-dimensionaler \mathbb{K} -Vektorraum, $\varphi: V \rightarrow V$ linear und λ ein Eigenwert von φ . Unter dem *verallgemeinerten Eigenraum* von φ zum Eigenwert λ verstehen wir den Teilraum

$$\tilde{E}_\lambda := \{v \in V \mid \exists N \in \mathbb{N} : (\varphi - \lambda \text{id}_V)^N(v) = 0\} = \bigcup_{N \geq 1} \ker((\varphi - \lambda \text{id}_V)^N).$$

Beachte, dass \tilde{E}_λ wegen der offensichtlichen Inklusionen $\ker((\varphi - \lambda \text{id}_V)^N) \subseteq \ker((\varphi - \lambda \text{id}_V)^{N+1})$ tatsächlich einen Teilraum von V bildet, der den Eigenraum zum Eigenwert λ enthält, $E_\lambda \subseteq \tilde{E}_\lambda$. Unter den verallgemeinerten Eigenräumen einer Matrix $A \in M_{n \times n}(\mathbb{K})$ verstehen wir die verallgemeinerten Eigenräume der damit assoziierten linearen Abbildung, $\mathbb{K}^n \rightarrow \mathbb{K}^n$, $x \mapsto Ax$.

VI.3.2. BEISPIEL. Die reelle Matrix

$$A = \begin{pmatrix} 5 & 1 & 0 \\ 0 & 5 & 1 \\ 0 & 0 & 5 \end{pmatrix}$$

hat charakteristisches Polynom $p = (5 - z)^3$ und daher nur einen Eigenwert, $\lambda = 5$, mit algebraischer Vielfachheit 3. Der entsprechende Eigenraum, $E_5 = \ker(A - 5I) = \langle e_1 \rangle$, ist ein-dimensional. Da

$$A - 5I_3 = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}, \quad (A - 5I_3)^2 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad (A - 5I_3)^3 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix},$$

ist der verallgemeinerte Eigenraum, $\tilde{E}_5 = \ker((A - 5I)^3) = \mathbb{K}^3$, echt größer als E_5 .

VI.3.3. BEISPIEL. Die reelle Matrix

$$A = \begin{pmatrix} 3 & 1 & 0 & 0 \\ 3 & 3 & 0 & 0 \\ 0 & 3 & 1 & 0 \\ 0 & 0 & 5 & 5 \end{pmatrix}$$

hat charakteristisches Polynom $p = (3 - z)^4(5 - z)^2$, also zwei Eigenwerte, nämlich $\lambda_1 = 3$ mit algebraischer Vielfachheit 4 und $\lambda_2 = 5$ mit algebraischer Vielfachheit 2. Aus

$$A - 3I_5 = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 2 \\ 0 & 0 & 0 & 0 & 2 \end{pmatrix}, \quad (A - 3I_5)^2 = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 4 & 0 \\ 0 & 0 & 0 & 0 & 4 \end{pmatrix},$$

erhalten wir $E_3 = \ker(A - 3I) = \langle e_1, e_3 \rangle$ und $\tilde{E}_3 = \ker((A - 3I)^2) = \langle e_1, e_2, e_3, e_4 \rangle$, der verallgemeinerte Eigenraum ist daher echt größer als der Eigenraum. Ebenso,

$$A - 5I_5 = \begin{pmatrix} -2 & 1 & 0 & 0 & 0 \\ -2 & -2 & 0 & 0 & 0 \\ 0 & -2 & 1 & 0 & 0 \\ 0 & 0 & -2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}, \quad (A - 5I_5)^2 = \begin{pmatrix} 4 & * & 0 & 0 & 0 \\ 4 & 4 & * & 0 & 0 \\ 4 & 4 & * & 4 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix},$$

und daher $E_5 = \tilde{E}_5 = \ker(A - 5I) = \langle e_5, e_6 \rangle$, für diesen Eigenwert stimmt der verallgemeinerte Eigenraum also mit dem Eigenraum überein.

Wir untersuchen zunächst verallgemeinerte Eigenräume zum Eigenwert 0.

VI.3.4. LEMMA. Sei V ein n -dimensionaler Vektorraum über einem Körper \mathbb{K} und $\varphi: V \rightarrow V$ linear. Dann existiert $N \in \{0, \dots, n\}$, sodass

$$\ker(\varphi^N) = \ker(\varphi^{N+1}).$$

Es gilt dann für jedes $k \in \mathbb{N}$,

$$\ker(\varphi^N) = \ker(\varphi^{N+k}) \quad \text{und} \quad \text{img}(\varphi^N) = \text{img}(\varphi^{N+k}).$$

Darüber hinaus gilt

$$V = \ker(\varphi^N) \oplus \text{img}(\varphi^N),$$

und diese Zerlegung ist invariant unter φ , d.h.

$$\varphi(\ker(\varphi^N)) \subseteq \ker(\varphi^N) \quad \text{und} \quad \varphi(\text{img}(\varphi^N)) \subseteq \text{img}(\varphi^N).$$

Weiters ist die Einschränkung $\varphi|_{\ker(\varphi^N)}: \ker(\varphi^N) \rightarrow \ker(\varphi^N)$ triangulierbar mit 0 als einzigem Eigenwert, und die Einschränkung $\varphi|_{\text{img}(\varphi^N)}: \text{img}(\varphi^N) \rightarrow \text{img}(\varphi^N)$ ist invertierbar. Bezeichnet m die algebraische Vielfachheit des Eigenwerts 0 von φ , so ist

$$\dim(\ker(\varphi^N)) = m, \quad \ker(\varphi^N) = \ker(\varphi^m) \quad \text{und} \quad \text{img}(\varphi^N) = \text{img}(\varphi^m).$$

Schließlich existiert eine Basis B von V , sodass

$$[\varphi]_{BB} = \begin{pmatrix} A & 0 \\ 0 & C \end{pmatrix}, \quad A = \begin{pmatrix} 0 & * & \cdots & * \\ 0 & 0 & \ddots & \vdots \\ \vdots & \ddots & \ddots & * \\ 0 & \cdots & 0 & 0 \end{pmatrix} \in M_{m \times m}(\mathbb{K}), \quad C \in \text{GL}_{n-m}(\mathbb{K}).$$

BEWEIS. Betrachte die aufsteigende Kette von Teilräumen:

$$\{0\} = \ker(\text{id}_V) = \ker(\varphi^0) \subseteq \ker(\varphi^1) \subseteq \ker(\varphi^2) \subseteq \cdots \subseteq \ker(\varphi^n) \subseteq \ker(\varphi^{n+1}) \subseteq V.$$

Indirekt angenommen $\ker(\varphi^N) \neq \ker(\varphi^{N+1})$, für alle $N \in \{0, \dots, n\}$. Dann gilt

$$\dim \ker(\varphi^N) < \dim \ker(\varphi^{N+1}), \quad N = 0, 1, \dots, n,$$

wir erhalten

$$0 = \dim \ker \varphi^0 < \dim \ker \varphi^1 < \dim \ker \varphi^2 < \cdots < \dim \ker \varphi^n < \dim \ker \varphi^{n+1},$$

und somit $\dim \ker(\varphi^{n+1}) \geq n + 1$. Da dies $\dim \ker(\varphi^{n+1}) \leq \dim(V) = n$ widerspricht, muss also ein $N \in \{0, \dots, n\}$ mit $\ker(\varphi^N) = \ker(\varphi^{N+1})$ existieren.

Wir zeigen nun $\ker(\varphi^N) = \ker(\varphi^{N+k})$ mittels Induktion nach k . Der Induktionsanfang, $k = 0$, ist trivial. Für den Induktionsschritt sei nun $k \geq 1$. Da offensichtlich $\ker(\varphi^N) \subseteq \ker(\varphi^{N+k})$, genügt es die umgekehrte Inklusion,

$$\ker(\varphi^N) \supseteq \ker(\varphi^{N+k}),$$

zu zeigen. Sei dazu $v \in \ker(\varphi^{N+k})$. Dann gilt $\varphi^{N+1}(\varphi^{k-1}(v)) = \varphi^{N+k}(v) = 0$, also $\varphi^{k-1}(v) \in \ker(\varphi^{N+1}) = \ker(\varphi^N)$ und daher $v \in \ker(\varphi^{N+k-1}) = \ker(\varphi^N)$, wobei

wir im letzten Gleichheitszeichen, die Induktionsvoraussetzung verwendet haben. Damit ist $\ker(\varphi^N) = \ker(\varphi^{N+k})$ gezeigt. Insbesondere gilt

$$\dim(\ker(\varphi^N)) = \dim(\ker(\varphi^{N+k}))$$

und wegen der Dimensionsformel, $\dim(V) = \dim(\ker(\varphi^l)) + \dim(\operatorname{img}(\varphi^l))$, auch

$$\dim(\operatorname{img}(\varphi^N)) = \dim(\operatorname{img}(\varphi^{N+k}))$$

für alle $k \in \mathbb{N}$. Da $\operatorname{img}(\varphi^{N+k}) \subseteq \operatorname{img}(\varphi^N)$ folgt $\operatorname{img}(\varphi^N) = \operatorname{img}(\varphi^{N+k})$.

Wir zeigen nun

$$\ker(\varphi^N) \cap \operatorname{img}(\varphi^N) = \{0\}. \quad (\text{VI.21})$$

Sei dazu $v \in \ker(\varphi^N) \cap \operatorname{img}(\varphi^N)$. Dann existiert $w \in V$, sodass $v = \varphi^N(w)$. Somit $\varphi^{N+N}(w) = \varphi^N(v) = 0$, also $w \in \ker(\varphi^{N+N}) = \ker(\varphi^N)$. Dies bedeutet $v = \varphi^N(w) = 0$. Damit ist (VI.21) gezeigt. Zusammen mit der Dimensionsformel

$$\dim(V) = \dim(\ker(\varphi^N)) + \dim(\operatorname{img}(\varphi^N))$$

folgt, dass V direkte Summe der Teilräume $\ker(\varphi^N)$ und $\operatorname{img}(\varphi^N)$ ist. Die Invarianz der Zerlegung ist offensichtlich, denn $\varphi(\ker(\varphi^N)) \subseteq \ker(\varphi^{N-1}) \subseteq \ker(\varphi^N)$ und $\varphi(\operatorname{img}(\varphi^N)) = \operatorname{img}(\varphi^{N+1}) = \operatorname{img}(\varphi^N)$. Daraus folgt auch, dass die Einschränkung $\varphi|_{\operatorname{img}(\varphi^N)}: \operatorname{img}(\varphi^N) \rightarrow \operatorname{img}(\varphi^N)$ surjektiv, also invertierbar ist. Wähle eine Basis b_1, \dots, b_{k_1} von $\ker(\varphi)$, ergänze sie zu einer Basis $b_1, \dots, b_{k_1}, \dots, b_{k_2}$ von $\ker(\varphi^2)$, ergänze weiter zu einer Basis $b_1, \dots, b_{k_1}, \dots, b_{k_2}, \dots, b_{k_3}$ von $\ker(\varphi^3)$, u.s.w. bis wir bei einer Basis $\tilde{B} := (b_1, \dots, b_k)$ von $\ker(\varphi^N)$ angekommen sind, wobei $k := \dim(\ker(\varphi^N))$. Bezüglich dieser Basis hat $\varphi|_{\ker(\varphi^N)}$ folgende Blockdiagonalgestalt:

$$A := [\varphi|_{\ker(\varphi^N)}]_{\tilde{B}\tilde{B}} = \begin{pmatrix} 0 & * & \cdots & * \\ 0 & 0 & \ddots & \vdots \\ \vdots & \ddots & \ddots & * \\ 0 & \cdots & 0 & 0 \end{pmatrix} \in M_{k \times k}(\mathbb{K}). \quad (\text{VI.22})$$

Dies zeigt, dass die Einschränkung $\varphi|_{\ker(\varphi^N)}$ triangulierbar ist mit 0 als einzigem Eigenwert. Sei nun b_{k+1}, \dots, b_n eine Basis von $\operatorname{img}(\varphi^N)$, und betrachte die Basis $B := (b_1, \dots, b_n)$ von V . Da φ auf $\operatorname{img}(\varphi^N)$ invertierbar ist, gilt

$$[\varphi]_{BB} = \begin{pmatrix} A & 0 \\ 0 & C \end{pmatrix}$$

mit obigem A und $C \in \operatorname{GL}_{n-k}(\mathbb{K})$. Insbesondere ist 0 nicht Eigenwert von C . Nach Lemma VI.2.14(d) gilt $p_\varphi = p_A p_C$, also muss der Eigenwert 0 von A algebraische Vielfachheit m haben, und wir erhalten $k = m$, vgl. (VI.22). Daraus folgt nun auch $A^m = 0$, also $[\varphi^m]_{BB} = \begin{pmatrix} 0 & 0 \\ 0 & C^m \end{pmatrix}$ mit $C^m \in \operatorname{GL}_{n-m}(\mathbb{K})$. Daraus lesen wir $\ker(\varphi^N) = \ker(\varphi^m)$ und $\operatorname{img}(\varphi^N) = \operatorname{img}(\varphi^m)$ ab. Damit ist der Beweis des Lemmas vollständig. \square

Aus dem vorangehenden Lemma erhalten wir sofort eine analoge Aussage über verallgemeinerte Eigenräume zu beliebigen Eigenwerten:

VI.3.5. LEMMA. Sei V ein n -dimensionaler Vektorraum über einem Körper \mathbb{K} und $\varphi: V \rightarrow V$ linear. Weiters sei λ Eigenwert von φ mit algebraischer Vielfachheit m und verallgemeinerten Eigenraum \tilde{E}_λ . Dann gilt

$$\dim(\tilde{E}_\lambda) = m \quad \text{und} \quad \tilde{E}_\lambda = \ker((\varphi - \lambda \text{id}_V)^m).$$

Weiters existiert ein invarianter komplementärer Teilraum W , d.h.

$$V = \tilde{E}_\lambda \oplus W, \quad \varphi(\tilde{E}_\lambda) \subseteq \tilde{E}_\lambda, \quad \varphi(W) \subseteq W.$$

Darüber hinaus ist die Einschränkung $\varphi|_{\tilde{E}_\lambda}: \tilde{E}_\lambda \rightarrow \tilde{E}_\lambda$ triangulierbar mit λ als einzigen Eigenwert, und λ ist kein Eigenwert der Einschränkung $\varphi|_W: W \rightarrow W$, d.h.

$$\sigma(\varphi|_{\tilde{E}_\lambda}) = \{\lambda\} \quad \text{und} \quad \sigma(\varphi|_W) = \sigma(\varphi) \setminus \{\lambda\}.$$

Schließlich existiert eine Basis B von V , sodass

$$[\varphi]_{BB} = \begin{pmatrix} A & 0 \\ 0 & C \end{pmatrix}, \quad A = \begin{pmatrix} \lambda & * & \cdots & * \\ 0 & \lambda & \ddots & \vdots \\ \vdots & \ddots & \ddots & * \\ 0 & \cdots & 0 & \lambda \end{pmatrix} \in M_{m \times m}(\mathbb{K}), \quad (\text{VI.23})$$

und $C \in M_{(n-m) \times (n-m)}(\mathbb{K})$ mit $\sigma(C) = \sigma(\varphi) \setminus \{\lambda\}$.

BEWEIS. Beachte, dass 0 Eigenwert der linearen Abbildung

$$\varphi - \lambda \text{id}_V: V \rightarrow V$$

ist und algebraische Vielfachheit m besitzt. Wenden wir Lemma VI.3.4 auf diese lineare Abbildung an, so folgt $\dim(\tilde{E}_\lambda) = m$, $\tilde{E}_\lambda = \ker((\varphi - \lambda \text{id}_V)^m)$, und der Teilraum $W := \text{img}((\varphi - \lambda \text{id}_V)^m)$ bildet ein Komplement,

$$V = \tilde{E}_\lambda \oplus W.$$

Da beide Teilräume unter $\varphi - \lambda \text{id}_V$ invariant sind, sind sie auch invariant unter φ , d.h. $\varphi(\tilde{E}_\lambda) \subseteq \tilde{E}_\lambda$ und $\varphi(W) \subseteq W$. Nach Lemma ist $(\varphi - \lambda \text{id}_V)|_{\tilde{E}_\lambda}$ triangulierbar mit 0 als einzigen Eigenwert. Somit ist auch $\varphi|_{\tilde{E}_\lambda}$ triangulierbar mit λ als einzigen Eigenwert, insbesondere gilt $\sigma(\varphi|_{\tilde{E}_\lambda}) = \{\lambda\}$. Nach Lemma ist $(\varphi - \lambda \text{id}_V)|_W$ invertierbar, folglich ist λ kein Eigenwert von $\varphi|_W$, d.h. $\lambda \notin \sigma(\varphi|_W)$. Aus der Invarianz der Zerlegung folgt $\sigma(\varphi) = \sigma(\varphi|_{\tilde{E}_\lambda}) \cup \sigma(\varphi|_W)$, also $\sigma(\varphi|_W) = \sigma(\varphi) \setminus \{\lambda\}$. Sei nun B eine Basis von V wie in Lemma VI.3.4, d.h.

$$[\varphi - \lambda \text{id}_V]_{BB} = \begin{pmatrix} \tilde{A} & 0 \\ 0 & \tilde{C} \end{pmatrix}, \quad \tilde{A} = \begin{pmatrix} 0 & * & \cdots & * \\ 0 & 0 & \ddots & \vdots \\ \vdots & \ddots & \ddots & * \\ 0 & \cdots & 0 & 0 \end{pmatrix} \in M_{m \times m}(\mathbb{K}),$$

und $\tilde{C} \in \text{GL}_{n-m}(\mathbb{K})$. Dann hat $[\varphi]_{BB}$ die Form (VI.23), wobei $A = \tilde{A} + \lambda I_m$ und $C = \tilde{C} + \lambda I_{n-m}$. \square

VI.3.6. BEMERKUNG. Betrachte nochmals die Matrix A aus Beispiel VI.3.3. Sie hat $\lambda = 3$ als Eigenwert mit algebraischer Vielfachheit 4. Nach dem vorangehenden Lemma muss $\tilde{E}_3 = \ker((A - 3I)^4)$ gelten. Weiter oben haben wir gesehen, dass in diesem Fall sogar $\tilde{E}_3 = \ker((A - 3I)^2)$ gilt. I.A. ist daher die algebraische Vielfachheit, m , eines Eigenwerts λ , nicht die minimale Zahl mit $\tilde{E}_\lambda = \ker((A - \lambda I)^m)$.

VI.3.7. SATZ (Primärzerlegung). Sei V ein n -dimensionaler Vektorraum über einem Körper \mathbb{K} und $\varphi: V \rightarrow V$ linear. Weiters bezeichnen $\lambda_1, \dots, \lambda_k$ alle (paarweise verschiedenen) Eigenwerte, m_1, \dots, m_k ihre algebraischen Vielfachheiten und $\tilde{E}_{\lambda_1}, \dots, \tilde{E}_{\lambda_k}$ die entsprechenden verallgemeinerten Eigenräume. Dann gilt

$$\dim(\tilde{E}_{\lambda_i}) = m_i \quad \text{und} \quad \tilde{E}_{\lambda_i} = \ker((\varphi - \lambda_i \text{id}_V)^{m_i}), \quad (\text{VI.24})$$

für alle $i = 1, \dots, k$. Darüber hinaus existiert ein invarianter Teilraum W , sodass

$$V = \tilde{E}_{\lambda_1} \oplus \dots \oplus \tilde{E}_{\lambda_k} \oplus W, \quad \varphi(\tilde{E}_{\lambda_i}) \subseteq \tilde{E}_{\lambda_i}, \quad \varphi(W) \subseteq W. \quad (\text{VI.25})$$

Die Einschränkung $\varphi|_{\tilde{E}_{\lambda_i}}: \tilde{E}_{\lambda_i} \rightarrow \tilde{E}_{\lambda_i}$ ist triangulierbar mit λ_i als einzigen Eigenwert, und $\varphi|_W$ hat keine Eigenwerte, also

$$\sigma(\varphi|_{\tilde{E}_{\lambda_i}}) = \{\lambda_i\} \quad \text{und} \quad \sigma(\varphi|_W) = \emptyset.$$

Schließlich existiert eine Basis B von V , sodass

$$[\varphi]_{BB} = \begin{pmatrix} A_1 & 0 & \dots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & A_k & 0 \\ 0 & \dots & 0 & A \end{pmatrix}, \text{ wobei } A_i = \begin{pmatrix} \lambda_i & * & \dots & * \\ 0 & \lambda_i & \ddots & \vdots \\ \vdots & \ddots & \ddots & * \\ 0 & \dots & 0 & \lambda_i \end{pmatrix} \in M_{m_i \times m_i}(\mathbb{K})$$

und $A \in M_{m \times m}(\mathbb{K})$ keine Eigenwerte besitzt, $\sigma(A) = \emptyset$, und $m = n - m_1 - \dots - m_k$.

Zerfällt das charakteristische Polynom von φ in Linearfaktoren (etwa weil \mathbb{K} algebraisch abgeschlossen ist), dann ist $W = \{0\}$, $V = \tilde{E}_{\lambda_1} \oplus \dots \oplus \tilde{E}_{\lambda_k}$ und der letzte Block in der Matrixdarstellung, A , tritt nicht auf, d.h. $m = 0$.

BEWEIS. Wir gehen mittels Induktion nach der Anzahl der verschiedenen Eigenwerte vor. Der Induktionsanfang, $\sigma(\varphi) = \emptyset$, ist trivial. Für den Induktionsschritt sei nun λ ein Eigenwert von φ . Nach Lemma VI.3.5 existiert ein komplementärer Teilraum W' , sodass

$$V = \tilde{E}_{\lambda_1} \oplus W', \quad \varphi(\tilde{E}_{\lambda_1}) \subseteq \tilde{E}_{\lambda_1}, \quad \varphi(W') \subseteq W'. \quad (\text{VI.26})$$

Bezeichnet $\varphi' := \varphi|_{W'}: W' \rightarrow W'$ die Einschränkung, so gilt weiters

$$\sigma(\varphi|_{\tilde{E}_{\lambda_1}}) = \{\lambda_1\} \quad \text{und} \quad \sigma(\varphi') = \{\lambda_2, \dots, \lambda_n\}.$$

Nach Induktionsvoraussetzung existiert ein Teilraum W von W' , sodass

$$W' = \tilde{E}'_{\lambda_2} \oplus \cdots \oplus \tilde{E}'_{\lambda_k} \oplus W, \quad \varphi'(\tilde{E}'_{\lambda_i}) \subseteq \tilde{E}'_{\lambda_i}, \quad \varphi'(W) \subseteq W, \quad \sigma(\varphi'|_W) = \emptyset,$$

wobei \tilde{E}'_{λ_i} die verallgemeinerten Eigenräume von φ' bezeichnen. Offensichtlich ist $\tilde{E}'_{\lambda_i} \subseteq \tilde{E}_{\lambda_i}$, es gilt aber auch die umgekehrte Inklusion, $\tilde{E}_{\lambda_i} \subseteq \tilde{E}'_{\lambda_i}$, für jedes $i = 2, \dots, k$. Sei dazu $v \in \tilde{E}_{\lambda_i}$ und N so, dass $(\varphi - \lambda_i \text{id}_V)^N v = 0$. Bezeichnet $v = e + w'$ die eindeutige Zerlegung mit $e \in \tilde{E}_{\lambda_1}$ und $w' \in W'$, dann folgt $(\varphi - \lambda_i \text{id}_V)^N e = 0 = (\varphi - \lambda_i \text{id}_V)^N w'$ wegen der Invarianz der Zerlegung (VI.26). Da λ_i nicht Eigenwert von $\varphi|_{\tilde{E}_{\lambda_1}}$ ist, muss $e = 0$ gelten, und daher $v = w' \in \tilde{E}'_{\lambda_i}$. Dies zeigt $\tilde{E}'_{\lambda_i} = \tilde{E}_{\lambda_i}$, und daher die invariante Zerlegung (VI.25). Zerfällt das charakteristische Polynom in Linearfaktoren, dann muss $W = \{0\}$ gelten, da ja $\sigma(\varphi'|_W) = \emptyset$. Wählen wir Basen von \tilde{E}_{λ_i} wie in Lemma VI.3.5 und vereinigen diese mit einer Basis von W , so erhalten wir eine Basis B von V bezüglich der die Matrixdarstellung $[\varphi]_{BB}$ die gewünschte Form hat. \square

Daraus erhalten wir sofort:

VI.3.8. KOROLLAR. *Sei V ein endlich dimensionaler Vektorraum über einem Körper \mathbb{K} und $\varphi: V \rightarrow V$ linear. Weiters bezeichnen $\lambda_1, \dots, \lambda_k$ alle (paarweise verschiedenen) Eigenwerte und $\tilde{E}_{\lambda_1}, \dots, \tilde{E}_{\lambda_k}$ die entsprechenden verallgemeinerten Eigenräume. Dann sind äquivalent:*

- (a) *Das charakteristische Polynom p_φ zerfällt über \mathbb{K} in Linearfaktoren.*
- (b) $V = \tilde{E}_{\lambda_1} \oplus \cdots \oplus \tilde{E}_{\lambda_k}$.
- (c) $\dim(V) = \sum_{i=1}^k \dim(\tilde{E}_{\lambda_i})$.

VI.3.9. BEISPIEL. Wir wollen die Primärzerlegung der Matrix

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

bestimmen. Für das charakteristische Polynom erhalten wir

$$\begin{aligned} p &= \det \begin{pmatrix} 5-z & 1 & 1 & 0 & 0 \\ -4 & 1-z & -2 & 0 & 0 \\ 0 & 0 & 3-z & 0 & 0 \\ 0 & 0 & 6 & 6-z & 1 \\ 0 & 0 & -6 & -1 & 4-z \end{pmatrix} \\ &= \det \begin{pmatrix} 5-z & 1 & 1 \\ -4 & 1-z & -2 \\ 0 & 0 & 3-z \end{pmatrix} \det \begin{pmatrix} 6-z & 1 \\ -1 & 4-z \end{pmatrix} = (3-z)^3(5-z)^2, \end{aligned}$$

also $\sigma(A) = \{3, 5\}$. Dann ist

$$A - 5I = \begin{pmatrix} 0 & 1 & 1 & 0 & 0 \\ -4 & -4 & -2 & 0 & 0 \\ 0 & 0 & -2 & 0 & 0 \\ 0 & 0 & 6 & 1 & 1 \\ 0 & 0 & -6 & -1 & -1 \end{pmatrix}, \quad (A - 5I)^2 = \begin{pmatrix} -4 & -4 & -4 & 0 & 0 \\ 16 & 12 & 8 & 0 & 0 \\ 0 & 0 & 4 & 0 & 0 \\ 0 & 0 & -12 & 0 & 0 \\ 0 & 0 & 12 & 0 & 0 \end{pmatrix},$$

also

$$E_5 = \ker(A - 5I) = \left\langle \left(\begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \\ -1 \end{pmatrix} \right) \right\rangle, \quad \tilde{E}_5 = \ker((A - 5I)^2) = \left\langle \left(\begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \\ -1 \end{pmatrix} \right), \left(\begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} \right) \right\rangle,$$

siehe (VI.24). Analog

$$A - 3I = \begin{pmatrix} 2 & 1 & 1 & 0 & 0 \\ -4 & -2 & -2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 6 & 3 & 1 \\ 0 & 0 & -6 & -1 & 1 \end{pmatrix}, \quad (A - 3I)^2 = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 12 & 8 & 4 \\ 0 & 0 & -12 & -4 & 0 \end{pmatrix},$$

also

$$E_3 = \ker(A - 3I) = \left\langle \left(\begin{pmatrix} 1 \\ -2 \\ 0 \\ 0 \\ 0 \end{pmatrix} \right), \left(\begin{pmatrix} 0 \\ 1 \\ -1 \\ 3 \\ -3 \end{pmatrix} \right) \right\rangle,$$

und

$$\tilde{E}_3 = \ker((A - 3I)^2) = \left\langle \left(\begin{pmatrix} 1 \\ -2 \\ 0 \\ 0 \\ 0 \end{pmatrix} \right), \left(\begin{pmatrix} 0 \\ 1 \\ -1 \\ 3 \\ -3 \end{pmatrix} \right), \left(\begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} \right) \right\rangle.$$

Bezeichnen wir die gewonnene Basis von V mit

$$B = \left(\begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \\ -1 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ -2 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ -1 \\ 3 \\ -3 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} \right),$$

und die entsprechende Basiswechselmatrix mit

$$S = T_{EB} = \begin{pmatrix} 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & -2 & 1 & 0 \\ 0 & 0 & 0 & -1 & 0 \\ 1 & 0 & 0 & 3 & 0 \\ -1 & 1 & 0 & -3 & 0 \end{pmatrix}$$

dann gilt

$$[A]_{BB} = S^{-1}AS = \left(\begin{array}{cc|ccc} 5 & 1 & 0 & 0 & 0 \\ 0 & 5 & 0 & 0 & 0 \\ \hline 0 & 0 & 3 & 0 & 2 \\ 0 & 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 0 & 3 \end{array} \right),$$

die gesuchte Primärzerlegung von A .

VI.3.10. SATZ (Caley–Hamilton). Sei V ein endlich dimensionaler Vektorraum über einem Körper \mathbb{K} und $\varphi: V \rightarrow V$ eine lineare Abbildung mit charakteristischem Polynom p . Dann gilt $p(\varphi) = 0$.

BEWEIS. Wir werden dies nur für algebraisch abgeschlossene Körper zeigen, der allgemeine Fall lässt sich dann durch Übergang zum algebraischen Abschluss von \mathbb{K} beweisen. Sei also \mathbb{K} algebraisch abgeschlossen. Es bezeichnen $\lambda_1, \dots, \lambda_k$ alle verschiedenen Eigenwerte, m_1, \dots, m_k ihre algebraischen Vielfachheiten und \tilde{E}_{λ_i} die entsprechenden verallgemeinerten Eigenräume. Für das charakteristische Polynom gilt dann $p = (\lambda_1 - z)^{m_1} \cdots (\lambda_k - z)^{m_k}$, und daher

$$p(\varphi) = (\lambda_1 \operatorname{id}_V - \varphi)^{m_1} \cdots (\lambda_k \operatorname{id}_V - \varphi)^{m_k}.$$

Nach Satz VI.3.7 ist

$$V = \tilde{E}_{\lambda_1} \oplus \cdots \oplus \tilde{E}_{\lambda_k} \quad \text{und} \quad \varphi(\tilde{E}_{\lambda_i}) \subseteq \tilde{E}_{\lambda_i}.$$

Schränken wir $p(\varphi): V \rightarrow V$ auf den Teilraum $\tilde{E}_{\lambda_k} \subseteq V$ ein, erhalten wir

$$p(\varphi)|_{\tilde{E}_{\lambda_k}} = (\lambda_1 \operatorname{id}_V - \varphi)^{m_1} \cdots \underbrace{(\lambda_k \operatorname{id}_V - \varphi)^{m_k}}_{=0}|_{\tilde{E}_{\lambda_k}} = 0,$$

siehe (VI.24). Da die Faktoren $(\lambda_i \operatorname{id}_V - \varphi)^{m_i}$ miteinander kommutieren, gilt auch

$$p(\varphi)|_{\tilde{E}_{\lambda_i}} = (\lambda_1 \operatorname{id}_V - \varphi)^{m_1} \cdots \underbrace{(\lambda_i \operatorname{id}_V - \varphi)^{m_i}}_{=0}|_{\tilde{E}_{\lambda_i}} = 0,$$

und wir erhalten analog

$$p(\varphi)|_{\tilde{E}_{\lambda_i}} = 0, \quad i = 1, \dots, k.$$

Nachdem die Teilräume $\tilde{E}_{\lambda_1}, \dots, \tilde{E}_{\lambda_k}$ ganz V aufspannen, folgt $p(\varphi) = 0$. \square

VI.3.11. BEMERKUNG. Ist $A \in M_{n \times n}(\mathbb{K})$ und bezeichnet p das charakteristische Polynom von A , dann gilt $p(A) = 0$. Dies folgt sofort aus Satz VI.3.10, siehe auch Aufgabe 60. Etwa hat die Matrix

$$A = \begin{pmatrix} 1 & -2 & -2 \\ 2 & 6 & 4 \\ -1 & -2 & 0 \end{pmatrix}$$

charakteristisches Polynom $p = -z^3 + 7z^2 - 16z + 12 = (2 - z)^2(3 - z)$ und es gilt tatsächlich

$$\begin{aligned} p(A) &= -A^3 + 7A^2 - 16A + 12I_3 \\ &= - \begin{pmatrix} -11 & -38 & -38 \\ 38 & 84 & 76 \\ -19 & -38 & -30 \end{pmatrix} + 7 \begin{pmatrix} -1 & -10 & -10 \\ 10 & 24 & 20 \\ -5 & -10 & -6 \end{pmatrix} - 16 \begin{pmatrix} 1 & -2 & -2 \\ 2 & 6 & 4 \\ -1 & -2 & 0 \end{pmatrix} + 12 \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} 11 & 38 & 38 \\ -38 & -84 & -76 \\ 19 & 38 & 30 \end{pmatrix} + \begin{pmatrix} -7 & -70 & -70 \\ 70 & 168 & 140 \\ -35 & -70 & -42 \end{pmatrix} + \begin{pmatrix} -16 & 32 & 32 \\ -32 & -96 & -64 \\ 16 & 32 & 0 \end{pmatrix} + \begin{pmatrix} 12 & 0 & 0 \\ 0 & 12 & 0 \\ 0 & 0 & 12 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}. \end{aligned}$$

VI.3.12. BEMERKUNG. Es scheint verlockend, den Beweis von Satz VI.3.10 wie folgt kürzer zu führen: Da $p(z) = \det(A - zI_n)$ ist $p(A) = \det(A - AI_n) = \det(A - A) = \det(0) = 0$. *Dies ist jedoch kein korrekter Beweis!* Das Problem liegt beim vermeintlichen Gleichheitszeichen, $p(A) = \det(A - AI_n)$, denn auf der linken Seite steht eine *Matrix*, $p(A)$, wohingegen auf der rechten Seite der *Skalar* $\det(A - AI_n)$ steht.

VI.3.13. DEFINITION (Nilpotente Abbildungen). Sei V ein endlich dimensionaler \mathbb{K} -Vektorraum. Eine lineare Abbildung $\varphi: V \rightarrow V$ wird *nilpotent* genannt, falls $N \in \mathbb{N}$ existiert, sodass $\varphi^N = 0$. Eine Matrix $A \in M_{n \times n}(\mathbb{K})$ wird nilpotent genannt, wenn die damit assoziierte lineare Abbildung $\mathbb{K}^n \rightarrow \mathbb{K}^n$, $x \mapsto Ax$, nilpotent ist. Dies ist genau dann der Fall, wenn $A^N = 0$, für ein $N \in \mathbb{N}$.

Aus den vorangehenden Betrachtungen folgt sofort:

VI.3.14. PROPOSITION. Sei V ein n -dimensionaler Vektorraum über einem Körper \mathbb{K} und $\varphi: V \rightarrow V$ eine lineare Abbildung. Dann sind äquivalent:

- (a) φ ist nilpotent.
- (b) $\tilde{E}_0 = V$.
- (c) $\varphi^n = 0$.
- (d) Es existieren ineinander geschachtelte Teilräume (Flagge)

$$\{0\} = V_0 \subseteq V_1 \subseteq V_2 \subseteq \cdots \subseteq V_r = V,$$

sodass $\varphi(V_i) \subseteq V_{i-1}$, für alle $i = 1, \dots, r$.

- (e) Es existiert eine Basis B von V , bezüglich der die Matrixdarstellung von φ strikte obere Dreiecksgestalt hat, d.h.

$$[\varphi]_{BB} = \begin{pmatrix} 0 & * & \cdots & * \\ 0 & 0 & \ddots & \vdots \\ \vdots & \ddots & \ddots & * \\ 0 & \cdots & 0 & 0 \end{pmatrix}$$

Ist \mathbb{K} algebraisch abgeschlossen, dann ist dies weiters äquivalent zu

- (f) $\sigma(\varphi) \subseteq \{0\}$.

Eine Folge (strikt) geschachtelter Teilräume wie in (d) ist durch

$$\{0\} = \ker(\varphi^0) \subsetneq \ker(\varphi^1) \subsetneq \ker(\varphi^2) \subsetneq \cdots \subsetneq \ker(\varphi^r) = V$$

gegeben, wobei r die kleinste Zahl mit $\varphi^r = 0$ bezeichnet (Nilpotenzindex). Eine Basis wie in (e) lässt sich gewinnen, indem wir mit einer Basis von $\ker(\varphi)$ beginnen, diese zu einer Basis von $\ker(\varphi^2)$ ergänzen, weiter zu einer Basis von $\ker(\varphi^3)$ ergänzen, u.s.w. bis wir schließlich bei einer Basis von V landen.

VI.3.15. KOROLLAR. Sei V ein endlich dimensionaler Vektorraum über einem Körper \mathbb{K} und $\varphi: V \rightarrow V$ eine lineare Abbildung deren charakteristisches Polynom in Linearfaktoren zerfällt. Dann existiert eine eindeutig bestimmte diagonalisierbare lineare Abbildung $\psi: V \rightarrow V$ und eine eindeutig bestimmte nilpotente

lineare Abbildung $\nu: V \rightarrow V$, sodass

$$\varphi = \psi + \nu \quad \text{und} \quad \psi\nu = \nu\psi.$$

Darüber hinaus haben φ und ψ das selbe charakteristische Polynom, also die gleichen Eigenwerte und algebraischen Vielfachheiten.

BEWEIS. Die Existenz von ψ und ν folgt aus der Primärzerlegung. Ist B eine Basis von V bezüglich der $[\varphi]_{BB}$ von der Form in Satz VI.3.7 ist, dann definieren wir $[\psi]_{BB}$ (und damit ψ) durch die Diagonaleinträge von $[\varphi]_{BB}$ und $\nu := \varphi - \psi$. Eine einfache Überlegung zeigt, dass diese linearen Abbildungen alle gewünschten Eigenschaften haben.

Es bleibt daher nur noch die Eindeutigkeit zu zeigen. Seien dazu $\psi: V \rightarrow V$ diagonalisierbar und $\nu: V \rightarrow V$ nilpotent, sodass $\varphi = \psi + \nu$ und $\psi\nu = \nu\psi$. Da die verallgemeinerten Eigenräume von φ ganz V aufspannen, genügt es $\psi|_{\tilde{E}_\lambda} = \lambda \text{id}_{\tilde{E}_\lambda}$ für jeden Eigenwert λ von φ zu zeigen, denn dadurch sind ψ und dann auch $\nu = \varphi - \psi$ eindeutig festgelegt. Da ψ und ν kommutieren gilt auch $\varphi\psi = \psi\varphi$ und $\varphi\nu = \nu\varphi$ und daher, siehe Aufgabe 61,

$$\psi(\tilde{E}_\lambda) \subseteq \tilde{E}_\lambda \quad \text{und} \quad \nu(\tilde{E}_\lambda) \subseteq \tilde{E}_\lambda.$$

Beachte, dass $\nu|_{\tilde{E}_\lambda}$ und $\varphi|_{\tilde{E}_\lambda} - \lambda \text{id}_{\tilde{E}_\lambda}$ beide nilpotent sind. Da sie kommutieren ist auch ihre Differenz, $\psi|_{\tilde{E}_\lambda} - \lambda \text{id}_{\tilde{E}_\lambda} = \varphi|_{\tilde{E}_\lambda} - \lambda \text{id}_{\tilde{E}_\lambda} - \nu|_{\tilde{E}_\lambda}$ nilpotent, siehe Aufgabe 62. Andererseits ist $\psi|_{\tilde{E}_\lambda} - \lambda \text{id}_{\tilde{E}_\lambda}$ auch diagonalisierbar, siehe Aufgabe 63. Somit ist $\psi|_{\tilde{E}_\lambda} - \lambda \text{id}_{\tilde{E}_\lambda}$ diagonalisierbar und nilpotent, also Null, siehe Aufgabe 64. Wir erhalten $\psi|_{\tilde{E}_\lambda} = \lambda \text{id}_{\tilde{E}_\lambda}$. \square

Nach Korollar VI.3.15, lässt sich jede Matrix $A \in M_{n \times n}(\mathbb{K})$, deren charakteristisches Polynom in Linearfaktoren zerfällt, in der Form $A = D + N$ schreiben, wobei $D \in M_{n \times n}(\mathbb{K})$ diagonalisierbar, $N \in M_{n \times n}(\mathbb{K})$ nilpotent, und $DN = ND$. Die Matrizen D und N sind durch diese Eigenschaften eindeutig bestimmt, die Eigenwerte von A und D , und deren algebraische Vielfachheiten, stimmen überein.

VI.3.16. SATZ. Sei V ein n -dimensionaler \mathbb{K} -Vektorraum und $\varphi: V \rightarrow V$ eine nilpotente lineare Abbildung. Dann existieren Vektoren $v_1, \dots, v_l \in V$ und Zahlen $r_1 \geq r_2 \geq \dots \geq r_l \geq 1$, sodass $\varphi^{r_i}(v_i) = 0$, für jedes $i = 1, \dots, l$, und so, dass

$$v_1, \varphi(v_1), \dots, \varphi^{r_1-1}(v_1), v_2, \varphi(v_2), \dots, \varphi^{r_2-1}(v_2), \dots, v_l, \varphi(v_l), \dots, \varphi^{r_l-1}(v_l)$$

eine Basis von V bildet. Bezeichnet B diese Basis in umgekehrter Reihenfolge, dann ist die Matrixdarstellung von der Form

$$[\varphi]_{BB} = \begin{pmatrix} 0 & \varepsilon_1 & 0 & \cdots & 0 \\ 0 & 0 & \varepsilon_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ \vdots & & \ddots & 0 & \varepsilon_{n-1} \\ 0 & \cdots & \cdots & 0 & 0 \end{pmatrix} \quad (\text{VI.27})$$

für gewisse $\varepsilon_i \in \{0, 1\}$. Es gilt in diesem Fall

$$n = r_1 + \cdots + r_l \quad \text{und} \quad l = \dim \ker \varphi = 1 + \#\{i \mid \varepsilon_i = 0\}.$$

Eine solche Basis kann durch Lösen linearer Gleichungssysteme algorithmisch konstruiert werden (siehe den Beweis unten).

BEWEIS. Sei r_1 die kleinste Zahl, sodass $\varphi^{r_1} = 0$ und $\varphi^{r_1-1} \neq 0$. Es existiert daher $v_1 \in V$, sodass $\varphi^{r_1}(v_1) = 0$. Wir zeigen zunächst, dass die Vektoren

$$v_1, \varphi(v_1), \varphi^2(v_1), \dots, \varphi^{r_1-1}(v_1) \quad (\text{VI.28})$$

linear unabhängig in V sind. Seien dazu $\lambda_i \in \mathbb{K}$ mit

$$\lambda_0 v_1 + \lambda_1 \varphi^1(v_1) + \lambda_2 \varphi^2(v_1) + \cdots + \lambda_{r_1-1} \varphi^{r_1-1}(v_1) = 0.$$

Da $\varphi^{r_1}(v_1) = 0$, erhalten wir durch sukzessives Anwenden von φ , das System

$$\begin{aligned} \lambda_0 v_1 + \lambda_1 \varphi^1(v_1) + \lambda_2 \varphi^2(v_1) + \cdots + \lambda_{r_1-1} \varphi^{r_1-1}(v_1) &= 0 \\ \lambda_0 \varphi^1(v_1) + \lambda_1 \varphi^2(v_1) + \cdots + \lambda_{r_1-2} \varphi^{r_1-1}(v_1) &= 0 \\ \lambda_0 \varphi^2(v_1) + \cdots + \lambda_{r_1-3} \varphi^{r_1-1}(v_1) &= 0 \\ &\vdots \\ \lambda_0 \varphi^{r_1-2}(v_1) + \lambda_1 \varphi^{r_1-1}(v_1) &= 0 \\ \lambda_0 \varphi^{r_1-1}(v_1) &= 0 \end{aligned}$$

Da $\varphi^{r_1-1}(v_1) \neq 0$ folgt daraus $\lambda_0 = \lambda_1 = \cdots = \lambda_{r_1-1} = 0$ in dem wir Gleichungssystem von unten nach oben inspizieren. Dies zeigt, dass (VI.28) linear unabhängig sind. Bezeichne den davon aufgespannten Teilraum mit

$$U := \langle v_1, \varphi(v_1), \varphi^2(v_1), \dots, \varphi^{r_1-1}(v_1) \rangle.$$

Beachte, dass U invariant unter φ ist, d.h. $\varphi(U) \subseteq U$. Die Abbildung φ induziert daher eine Abbildung auf dem Quotientenraum,

$$\bar{\varphi}: \bar{V} \rightarrow \bar{V}, \quad \bar{\varphi}([v]) = [\varphi(v)], \quad \bar{V} := V/U.$$

Beachte $\bar{\varphi}^{r_1} = 0$, denn $\bar{\varphi}^{r_1}([v]) = [\varphi^{r_1}(v)] = [0] = 0$, für jedes $v \in V$. Insbesondere ist auch $\bar{\varphi}: \bar{V} \rightarrow \bar{V}$ nilpotent. Mittels Induktion nach der Dimension von V dürfen wir daher annehmen, dass Zahlen $r_1 \geq r_2 \geq \cdots \geq r_l \geq 1$ und Vektoren $\bar{v}_2, \dots, \bar{v}_l \in \bar{V}$ existieren, sodass $\bar{\varphi}^{r_i}(\bar{v}_i) = 0$ für jedes $i = 2, \dots, l$, und so, dass

$$\bar{v}_2, \bar{\varphi}(\bar{v}_2), \dots, \bar{\varphi}^{r_2-1}(\bar{v}_2), \dots, \bar{v}_l, \bar{\varphi}(\bar{v}_l), \dots, \bar{\varphi}^{r_l-1}(\bar{v}_l) \quad (\text{VI.29})$$

eine Basis von \bar{V} bildet. Wähle Repräsentanten $\tilde{v}_2, \dots, \tilde{v}_l \in V$, d.h. $\bar{v}_i = [\tilde{v}_i]$. Es folgt $[\varphi^{r_i}(\tilde{v}_i)] = \bar{\varphi}^{r_i}(\bar{v}_i) = 0$, also $\varphi^{r_i}(\tilde{v}_i) \in U$. Es existieren daher $\mu_{i,j} \in \mathbb{K}$, sodass

$$\varphi^{r_i}(\tilde{v}_i) = \sum_{j=0}^{r_1-1} \mu_{i,j} \varphi^j(v_1), \quad i = 2, \dots, l. \quad (\text{VI.30})$$

Anwenden von $\varphi^{r_1-r_i}$ liefert, unter Verwendung von $\varphi^{r_1} = 0$,

$$\begin{aligned} 0 = \varphi^{r_1}(\tilde{v}_i) &= \varphi^{r_1-r_i}(\varphi^{r_i}(\tilde{v}_i)) = \sum_{j=0}^{r_1-1} \mu_{i,j} \varphi^{r_1-r_i+j}(v_1) \\ &= \sum_{j=r_1-r_i}^{r_1-1+r_1-r_i} \mu_{i,j-r_1+r_i} \varphi^j(v_1) = \sum_{j=r_1-r_i}^{r_1-1} \mu_{i,j-r_1+r_i} \varphi^j(v_1) \end{aligned}$$

Aus der linearen Unabhängigkeit des Systems (VI.28) folgt $\mu_{i,0} = \mu_{i,1} = \dots = \mu_{i,r_i-1} = 0$. Zusammen mit (VI.30) erhalten wir also

$$\varphi^{r_i}(\tilde{v}_i) = \sum_{j=r_i}^{r_1-1} \mu_{i,j} \varphi^j(v_1), \quad i = 2, \dots, l.$$

Setzen wir

$$v_i := \tilde{v}_i - \sum_{j=r_i}^{r_1-1} \mu_{i,j} \varphi^{j-r_i}(v_1),$$

dann gilt somit

$$\varphi^{r_i}(v_i) = 0 \quad \text{und} \quad \bar{v}_i = [v_i], \quad i = 2, \dots, l.$$

Beachte, dass

$$v_2, \varphi(v_2), \dots, \varphi^{r_2-1}(v_2), \dots, v_l, \varphi(v_l), \dots, \varphi^{r_l-1}(v_l) \quad (\text{VI.31})$$

unter der kanonischen Projektion $V \rightarrow \bar{V}$, $v \mapsto [v]$, bijektiv auf die Basis (VI.29) abgebildet wird. Somit ist (VI.31) linear unabhängig in V und bildet die Basis eines zu U komplementären Teilraums. Da $v_1, \varphi(v_1), \dots, \varphi^{r_1-1}(v_1)$ eine Basis von U ist, bildet

$$v_1, \varphi(v_1), \dots, \varphi^{r_1-1}(v_1), v_2, \varphi(v_2), \dots, \varphi^{r_2-1}(v_2), \dots, v_l, \varphi(v_l), \dots, \varphi^{r_l-1}(v_l)$$

also eine Basis von V . Damit ist der erste Teil des Satzes gezeigt. Die verbleibenden Aussagen sind nun offensichtlich. \square

Für nilpotente Matrizen $A \in M_{n \times n}(\mathbb{K})$, besagt der vorangehende Satz, dass eine invertierbare Matrix $S \in \text{GL}_n(\mathbb{K})$ und Skalare $\varepsilon_i \in \{0, 1\}$ existieren, sodass $S^{-1}AS$ die Form (VI.27) hat.

VI.3.17. BEISPIEL. Die Matrix

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

ist nilpotent, denn

$$A^2 = \begin{pmatrix} 2 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ -4 & 0 & -2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 & 0 & 0 \\ -6 & 0 & -3 & 0 & 0 & 0 \end{pmatrix}, \quad A^3 = \begin{pmatrix} 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -3 & 0 & 0 \end{pmatrix}, \quad A^4 = 0.$$

Wir wollen eine invertierbare Matrix S finden, sodass $S^{-1}AS$ die Gestalt (VI.27) hat. Wir gehen wie im Beweis des vorangehenden Satzes vor. Da $A^3 \neq 0$ und $A^4 = 0$ ist $r_1 = 4$. Es ist v_1 so zu wählen, dass $A^3v_1 \neq 0$. Wir verwenden $v_1 = e_4$ und erhalten $A^4v_1 = 0$ sowie

$$v_1 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \quad Av_1 = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 2 \\ 0 \end{pmatrix}, \quad A^2v_1 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \\ -1 \\ 0 \end{pmatrix}, \quad A^3v_1 = \begin{pmatrix} 1 \\ 0 \\ -2 \\ 0 \\ 0 \\ -3 \end{pmatrix}.$$

Nach dem Beweis des Satzes oben, sind diese Vektoren linear unabhängig und spannen daher einen 4-dimensionalen Teilraum U auf. Um den Nilpotenzindex der induzierten Abbildung auf dem Quotientenraum \mathbb{K}^6/U zu bestimmen, beobachten wir, dass $\text{img}(A^2)$ zur Gänze in U liegt, aber $\text{img}(A) \not\subseteq U$, somit $r_2 = 2$. Es ist daher \tilde{v}_2 so zu bestimmen, dass $A\tilde{v}_2 \notin U$. Wir entscheiden uns für $\tilde{v}_2 = e_1$ und erhalten

$$\tilde{v}_2 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \quad A\tilde{v}_2 = \begin{pmatrix} 0 \\ 2 \\ 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \quad A^2\tilde{v}_2 = \begin{pmatrix} 2 \\ 0 \\ -4 \\ 0 \\ 0 \\ -6 \end{pmatrix} = 2A^3v_1 \in U.$$

Setzen wir $v_2 := \tilde{v}_2 - 2Av_1$, dann gilt $A^2v_2 = 0$ und

$$v_2 = \begin{pmatrix} 1 \\ 0 \\ -2 \\ 0 \\ -4 \\ 0 \end{pmatrix}, \quad Av_2 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 3 \\ 0 \\ 0 \end{pmatrix}.$$

Wir erhalten eine Basis B und die entsprechende Basiswechselmatrix S ,

$$\underbrace{\left(\begin{pmatrix} 1 \\ 0 \\ -2 \\ 0 \\ 0 \\ -3 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \\ -1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 2 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 3 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ -2 \\ 0 \\ -4 \\ 0 \end{pmatrix} \right)}_B, \quad S = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ -2 & 0 & 1 & 0 & 0 & -2 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & -1 & 2 & 0 & 3 & -4 \\ -3 & 0 & 0 & 0 & 0 & 0 \end{pmatrix},$$

mit der gewünschten Eigenschaft:

$$S^{-1}AS = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

VI.3.18. DEFINITION (Jordanblock). Unter dem Jordanblock der Größe $m \in \mathbb{N}$ mit Eigenwert $\lambda \in \mathbb{K}$ verstehen wir die Matrix

$$J_m(\lambda) := \begin{pmatrix} \lambda & 1 & & \\ & \lambda & \ddots & \\ & & \ddots & 1 \\ & & & \lambda \end{pmatrix} \in M_{m \times m}(\mathbb{K}).$$

VI.3.19. SATZ (Jordan Zerlegung). Sei V ein n -dimensionaler Vektorraum über einem Körper \mathbb{K} und $\varphi: V \rightarrow V$ eine lineare Abbildung, deren charakteristisches Polynom über \mathbb{K} in Linearfaktoren zerfällt. Weiters bezeichnen $\lambda_1, \dots, \lambda_k$ alle (paarweise verschiedenen) Eigenwerte von φ und m_1, \dots, m_k ihre algebraische Vielfachheiten und n_1, \dots, n_k ihre geometrischen Vielfachheiten. Dann existiert eine Basis B von V , sodass die Matrixdarstellung von φ folgende Blockdiagonalgestalt besitzt:

$$[\varphi]_{BB} = \begin{pmatrix} J_{m_{1,1}}(\lambda_1) & & & & & & & & & & \\ & \ddots & & & & & & & & & \\ & & J_{m_{1,n_1}}(\lambda_1) & & & & & & & & \\ \hline & & & \ddots & & & & & & & \\ & & & & J_{m_{k,1}}(\lambda_k) & & & & & & \\ & & & & & \ddots & & & & & \\ & & & & & & J_{m_{k,n_k}}(\lambda_k) & & & & \end{pmatrix}.$$

Dabei sind $m_{i,j} \in \mathbb{N}$ und $m_{i,1} + \dots + m_{i,n_i} = m_i$, für jedes $i = 1, \dots, k$. Für die Anzahl der Jordanblöcke zum Eigenwert λ_i und Größe mindestens r gilt

$$\#\{j \mid m_{i,j} \geq r\} = \dim \ker((\varphi - \lambda_i \text{id}_V)^r) - \dim \ker((\varphi - \lambda_i \text{id}_V)^{r-1}).$$

Insbesondere ist diese Matrixdarstellung von φ , bis auf die Reihenfolge der Jordanblöcke, eindeutig bestimmt. Sie wird die Jordan'sche Normalform von φ genannt.

BEWEIS. Bezeichnen $\tilde{E}_{\lambda_i} = \ker((\varphi - \lambda_i \text{id}_V)^{m_i})$ die verallgemeinerten Eigenräume von φ , so haben wir nach Satz VI.3.7 eine invariante Zerlegung,

$$V = \tilde{E}_{\lambda_1} \oplus \dots \oplus \tilde{E}_{\lambda_k}, \quad \varphi(\tilde{E}_{\lambda_i}) \subseteq \tilde{E}_{\lambda_i},$$

und die Einschränkung $\varphi|_{\tilde{E}_{\lambda_i}}$ ist triangulierbar mit einzigem Eigenwert λ_i . Somit ist die Einschränkung $\varphi|_{\tilde{E}_{\lambda_i}} - \lambda_i \text{id}_{\tilde{E}_{\lambda_i}} : \tilde{E}_{\lambda_i} \rightarrow \tilde{E}_{\lambda_i}$ nilpotent. Nach Satz VI.3.16 existiert daher eine Basis B_i von \tilde{E}_{λ_i} und Zahlen $\varepsilon_{i,j} \in \{0, 1\}$, sodass

$$[\varphi|_{\tilde{E}_{\lambda_i}} - \lambda_i \text{id}_{\tilde{E}_{\lambda_i}}]_{B_i B_i} = \begin{pmatrix} 0 & \varepsilon_{i,1} & & \\ & 0 & \ddots & \\ & & \ddots & \varepsilon_{i,m_i-1} \\ & & & 0 \end{pmatrix}.$$

Folglich:

$$[\varphi|_{\tilde{E}_{\lambda_i}}]_{B_i B_i} = \begin{pmatrix} \lambda_i & \varepsilon_{i,1} & & \\ & \lambda_i & \ddots & \\ & & \ddots & \varepsilon_{i,m_i-1} \\ & & & \lambda_i \end{pmatrix}.$$

Beachte,

$$\begin{aligned} n_i &= \dim(E_{\lambda_i}) = \dim(\ker(\varphi - \lambda_i \text{id}_V)) \\ &= \dim(\ker(\varphi|_{\tilde{E}_{\lambda_i}} - \lambda_i \text{id}_{\tilde{E}_{\lambda_i}})) = 1 + \#\{j \mid \varepsilon_{i,j} = 0\}. \end{aligned}$$

Es existieren daher Zahlen $m_{i,1}, \dots, m_{i,n_i} \in \mathbb{N}$, sodass $m_{i,1} + \dots + m_{i,n_i} = m_i$ und

$$[\varphi|_{\tilde{E}_{\lambda_i}}]_{B_i B_i} = \underbrace{\begin{pmatrix} J_{m_{i,1}}(\lambda_i) & & \\ & \ddots & \\ & & J_{m_{i,n_i}}(\lambda_i) \end{pmatrix}}_{m_i \times m_i}, \quad J_{m_{i,j}}(\lambda_i) = \underbrace{\begin{pmatrix} \lambda_i & 1 & & \\ & \lambda_i & \ddots & \\ & & \ddots & 1 \\ & & & \lambda_i \end{pmatrix}}_{m_{i,j} \times m_{i,j}}.$$

Die Vereinigung $B := B_1, \dots, B_k$ bildet daher eine Basis von V , bezüglich der φ die gewünschte Form hat. Daraus erhalten wir sofort

$$\begin{aligned} \dim(\ker(\varphi - \lambda_i \text{id}_V)^r) &= \dim(\ker(\varphi|_{\tilde{E}_{\lambda_i}} - \lambda_i \text{id}_{\tilde{E}_{\lambda_i}})^r) \\ &= \dim \ker \begin{pmatrix} J_{m_{i,1}}^{(0)} & & \\ & \ddots & \\ & & J_{m_{i,n_i}}^{(0)} \end{pmatrix}^r = \#\{j \mid m_{i,j} \geq 1\} + \dots + \#\{j \mid m_{i,j} \geq r\}, \end{aligned}$$

und daher $\#\{j \mid m_{i,j} \geq r\} = \dim \ker((\varphi - \lambda_i \text{id}_V)^r) - \dim \ker((\varphi - \lambda_i \text{id}_V)^{r-1})$. \square

Für Matrizen $A \in M_{n \times n}(\mathbb{K})$, deren charakteristisches Polynom über \mathbb{K} in Linearfaktoren zerfällt, besagt der vorangehende Satz gerade, dass eine invertierbare Matrix $S \in \text{GL}_n(\mathbb{K})$ existiert, sodass $S^{-1}AS$ wie die Matrix in Satz VI.3.19

aussieht. Eine mögliche Jordan'sche Normalform ist etwa:

$$\left(\begin{array}{ccc|ccc} 3 & 1 & & & & \\ & 3 & 1 & & & \\ & & 3 & 1 & & \\ \hline & & & 3 & 1 & \\ & & & & 3 & \\ & & & & & 3 \\ & & & & & & 5 & 1 \\ & & & & & & & 5 \\ & & & & & & & & 5 \end{array} \right)$$

wobei, wie bisher, alle nicht spezifizierten Eintragungen Null sind.

VI.3.20. KOROLLAR. *Zwei Matrizen $A, B \in M_{n \times n}(\mathbb{K})$, deren charakteristische Polynome in Linearfaktoren zerfallen, sind genau dann ähnlich, wenn ihre Jordan'schen Normalformen, bis auf die Reihenfolge der Jordanblöcke, gleich sind. Dies ist genau dann der Fall, wenn A und B die selben Eigenwerte, $\lambda_1, \dots, \lambda_k$, haben und für jedes $i = 1, \dots, k$ und jedes r die folgende Relation gilt:*

$$\dim \ker((A - \lambda_i I_n)^r) = \dim \ker((B - \lambda_i I_n)^r).$$

VI.3.21. BEMERKUNG. Durch die geometrischen und algebraischen Vielfachheiten der Eigenwerte ist die Jordan'sche Normalform noch nicht eindeutig festgelegt. Etwa haben die beiden Matrizen

$$A = \begin{pmatrix} 7 & 1 & & & & \\ & 7 & 0 & & & \\ & & 7 & 1 & & \\ & & & 7 & 1 & \\ & & & & 7 & 1 \\ & & & & & 7 \end{pmatrix} \quad \text{und} \quad B = \begin{pmatrix} 7 & 1 & & & & \\ & 7 & 1 & & & \\ & & 7 & 0 & & \\ & & & 7 & 1 & \\ & & & & 7 & 1 \\ & & & & & 7 \end{pmatrix}$$

beide nur einen Eigenwert, nämlich $\lambda = 7$ mit algebraischer Vielfachheit 6 und geometrischer Vielfachheit 2. Die beiden Matrizen sind jedoch nicht ähnlich, da ihre Jordan'schen Normalformen verschieden sind, es gibt daher *keine* invertierbare Matrix $S \in \text{GL}_6(\mathbb{K})$, sodass $S^{-1}AS = B$. Dass A und B nicht ähnlich sein können, folgt auch direkt aus $(B - 7I)^3 = 0$ und $(A - 7I)^3 \neq 0$.

VI.3.22. BEISPIEL. Wir wollen die Jordan'sche Normalform der Matrix

$$A = \begin{pmatrix} 7 & 1 & 0 & -1 & 0 & 0 & -15 \\ 0 & 7 & 0 & 0 & -1 & -2 & 0 \\ 0 & 0 & 7 & 1 & 0 & 0 & 7 \\ 0 & 0 & 0 & 7 & 1 & 2 & 0 \\ 0 & 0 & 0 & 0 & 7 & 0 & -2 \\ 0 & 0 & 0 & 0 & 0 & 7 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 7 \end{pmatrix}$$

bestimmen. Wir setzen $N := A - 7I_n$ und berechnen

$$N = \begin{pmatrix} 0 & 1 & 0 & -1 & 0 & 0 & -15 \\ 0 & 0 & 0 & 0 & -1 & -2 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 7 \\ 0 & 0 & 0 & 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -2 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}, \quad N^2 = \begin{pmatrix} 0 & 0 & 0 & 0 & -2 & -4 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}, \quad N^3 = 0.$$

Somit $\dim \ker(N) = 3$, $\dim \ker(N^2) = 6$ und $\dim \ker(N^r) = 0$, für $r \geq 3$. Nach Satz VI.3.19 hat die Jordan'sche Normalform von A daher 3 Jordanblöcke der Größen 3, 2, 2, und muss somit folgende Gestalt haben:

$$J = \left(\begin{array}{ccc|cc} 7 & 1 & & & \\ & 7 & 1 & & \\ & & 7 & & \\ \hline & & & 7 & 1 \\ & & & & 7 \\ & & & \hline & & & & 7 & 1 \\ & & & & & 7 \end{array} \right) \quad (\text{VI.32})$$

Um auch eine Matrix S mit $S^{-1}AS = J$ zu bestimmen gehen wir wie im Beweis von Satz VI.3.16 vor. Da $N^2 \neq 0$ und $N^3 = 0$, ist $r_1 = 3$. Es daher v_1 so zu wählen, dass $N^2v_1 \neq 0$. Wir verwenden $v_1 = e_5$ und erhalten:

$$v_1 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \quad Nv_1 = \begin{pmatrix} 0 \\ -1 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \quad N^2v_1 = \begin{pmatrix} -2 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \quad N^3v_1 = 0.$$

Da $\text{img}(N^2) \subseteq U_1 := \langle v_1, Nv_1, N^2v_1 \rangle$ aber $\text{img}(N) \not\subseteq U_1$ ist $r_2 = 2$. Es daher v_2 so zu wählen, dass $Nv_2 \notin U_1$ und $N^2v_2 = 0$. Wir verwenden $v_2 = e_2$ und erhalten

$$v_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \quad Nv_2 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \quad N^2v_2 = 0.$$

Da $\text{img}(N) \not\subseteq \langle v_1, Nv_1, N^2v_1, v_2, Nv_2 \rangle$, ist auch $r_3 = 2$. Es ist daher v_3 so zu wählen, dass $Nv_3 \notin \langle v_1, Nv_1, N^2v_1, v_2, Nv_2 \rangle$ und $N^2v_3 = 0$. Wir verwenden $v_3 = e_7$ und erhalten

$$v_3 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}, \quad Nv_3 = \begin{pmatrix} -15 \\ 0 \\ 7 \\ 0 \\ -2 \\ 1 \\ 0 \end{pmatrix}, \quad N^2v_3 = 0.$$

Insgesamt erhalten wir folgende Basis und Transformationsmatrix:

$$\underbrace{\left(\begin{pmatrix} -2 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ -1 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} -15 \\ 0 \\ 7 \\ 0 \\ -2 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} \right)}_B \quad S = \begin{pmatrix} -2 & 0 & 0 & 1 & 0 & -15 & 0 \\ 0 & -1 & 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 7 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & -2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

Nach Konstruktion gilt $S^{-1}AS = J$, siehe (VI.32).

VI.3.23. BEISPIEL. Betrachte die Folge $0, 0, 1, 6, 24, 80, 240, 672 \dots$. Genauer:

$$x_0 = 0, \quad x_1 = 0, \quad x_2 = 1, \quad x_n = 6x_{n-1} - 12x_{n-2} + 8x_{n-3}, \quad n \geq 3.$$

Wir wollen eine explizite Formel für x_n herleiten. Setzen wir

$$v_n := \begin{pmatrix} x_n \\ x_{n+1} \\ x_{n+2} \end{pmatrix}, \quad A := \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 8 & -12 & 6 \end{pmatrix},$$

dann gilt $v_{n+1} = Av_n$ und daher $v_n = A^n v_0$. Da das gesuchte x_n den ersten Eintrag von v_n bildet, genügt es also A^n zu berechnen. Das charakteristische Polynom von A ist $p = (2 - z)^3$, die Matrix hat also nur einen Eigenwert, nämlich 2. Es gilt

$$A - 2I = \begin{pmatrix} -2 & 1 & 0 \\ 0 & -2 & 1 \\ 8 & -12 & 4 \end{pmatrix}, \quad (A - 2I)^2 = \begin{pmatrix} 4 & -4 & 1 \\ 8 & -8 & 2 \\ 16 & -16 & 4 \end{pmatrix}, \quad (A - 2I)^3 = 0,$$

und daher

$$b_1 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \quad (A - 2I)b_1 = \begin{pmatrix} 0 \\ 1 \\ 4 \end{pmatrix}, \quad (A - 2I)^2 b_1 = \begin{pmatrix} 1 \\ 2 \\ 4 \end{pmatrix}$$

Daraus lesen wir die Jordan'sche Normalform von A ab:

$$S^{-1}AS = \begin{pmatrix} 2 & 1 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{pmatrix} \quad \text{wobei} \quad S = \begin{pmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 4 & 4 & 1 \end{pmatrix}, \quad S^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 4 & -4 & 1 \end{pmatrix}.$$

Damit können wir die Potenzen von A bestimmen, siehe Aufgabe 67,

$$A^n = S \begin{pmatrix} 2 & 1 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{pmatrix}^n S^{-1} = S \begin{pmatrix} 2^n & n2^{n-1} & \binom{n}{2}2^{n-2} \\ 0 & 2^n & n2^{n-1} \\ 0 & 0 & 2^n \end{pmatrix} S^{-1}$$

Somit

$$v_n = A^n v_0 = S \begin{pmatrix} 2^n & n2^{n-1} & \binom{n}{2}2^{n-2} \\ 0 & 2^n & n2^{n-1} \\ 0 & 0 & 2^n \end{pmatrix} \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = S \begin{pmatrix} \binom{n}{2}2^{n-2} \\ n2^{n-1} \\ 2^n \end{pmatrix} = \begin{pmatrix} \binom{n}{2}2^{n-2} \\ \binom{n+1}{2}2^{n-1} \\ \binom{n+2}{2}2^n \end{pmatrix}.$$

Der erste Eintrag liefert die gesuchte explizite Formel für x_n ,

$$x_n = \binom{n}{2}2^{n-2} = \frac{n(n-1)}{2}2^{n-2} = n(n-1)2^{n-3}.$$

VI.4. Reelle Normalformen. Wir wollen hier die Primärzerlegung des vorangehenden Abschnitts verfeinern und anschließend die reelle Jordan'sche Normalform besprechen. Wir beginnen mit der Primfaktorzerlegung von Polynomen.

VI.4.1. DEFINITION (Teiler). Seien $p, q \in \mathbb{K}[z]$ zwei Polynome. Wir sagen p teilt q , und schreiben $p|q$, falls ein Polynom $s \in \mathbb{K}[z]$ existiert, sodass $q = sp$. Jedes solche Polynom p wird als Teiler von q bezeichnet.

Jedes nicht-triviale konstante Polynom $0 \neq c \in \mathbb{K}[z]$ teilt jedes andere Polynom $p \in \mathbb{K}[z]$, denn $p = c(c^{-1}p)$, wobei $c^{-1}p \in \mathbb{K}[z]$. Auch gilt $cp|p$ für jedes Polynom p und jede Konstante $0 \neq c \in \mathbb{K}$, denn $p = c^{-1}(cp)$. Solche Teiler, d.h. konstante Polynome und konstante Vielfache von p , werden als *triviale Teiler* von p bezeichnet.

VI.4.2. DEFINITION (Irreduzible Polynome). Ein Polynom p mit $\deg(p) \geq 1$ wird *irreduzibel* genannt, wenn es neben den trivialen Teilern (konstante Polynome und konstante Vielfache von p) keine weiteren Teiler besitzt.

VI.4.3. BEISPIEL. Jedes Polynom ersten Grades ist irreduzibel, denn für den Grad gilt $\deg(p_1 p_2) = \deg(p_1) + \deg(p_2)$. Das Polynom $p = z^2 + 1 \in \mathbb{R}[z]$ ist nicht irreduzibel, denn es zerfällt über \mathbb{R} nicht in Linearfaktoren.

VI.4.4. DEFINITION (Teilerfremde Polynome). Polynome $p_1, \dots, p_k \in \mathbb{K}[z]$ werden *teilerfremd* oder *relativ prim* genannt, falls sie neben den konstanten Polynomen keine weiteren gemeinsamen Teiler besitzen, d.h. $q|p_1, \dots, q|p_k$ ist nur für konstante Polynome q möglich.

VI.4.5. LEMMA. Polynome $p_1, \dots, p_k \in \mathbb{K}[z]$ sind genau dann teilerfremd, wenn Polynome $q_1, \dots, q_k \in \mathbb{K}[z]$ existieren, sodass

$$q_1 p_1 + \dots + q_k p_k = 1.$$

BEWEIS. Gilt $q_1 p_1 + \dots + q_k p_k = 1$ und ist $s \in \mathbb{K}[z]$ ein gemeinsamer Teiler aller p_i , d.h. $s|p_1, \dots, s|p_k$, dann folgt $s|1$, also muss s konstant sein. Dies zeigt, dass p_1, \dots, p_k teilerfremde Polynome sind. Für die andere Implikation betrachte

$$J := \{q_1 p_1 + \dots + q_k p_k \mid q_i \in \mathbb{K}[z]\}.$$

Da wenigstens ein $p_i \neq 0$ gilt $J \neq \{0\}$. Sei nun $0 \neq a \in J$ mit minimalem Grad, d.h. $\deg(a) \leq \deg(b)$, für alle $0 \neq b \in J$. Wir zeigen nun, dass a jedes der Polynome p_i teilt, $i = 1, \dots, k$. Da $a \in J$ existieren $q_1, \dots, q_k \in \mathbb{K}[z]$ mit $a = q_1 p_1 + \dots + q_k p_k$. Nach Lemma VI.2.1 existieren $r, s \in \mathbb{K}[z]$ mit $p_1 = sa + r$ und $\deg(r) < \deg(a)$. Somit

$$r = (1 - sq_1)p_1 - \dots - sq_k p_k \in J,$$

und wegen der Minimalität von $\deg(a)$, also $r = 0$. Wir schließen $p_1 = sa$ und daher $a|p_1$. Völlig analog lässt sich $a|p_i$ für jedes $i = 1, \dots, k$ zeigen. Als gemeinsamer Teiler der teilerfremden Polynome p_1, \dots, p_k muss a konstant sein.

Wir erhalten somit $1 = a^{-1}a = a^{-1}(q_1p_1 + \cdots + q_kp_k) = \tilde{q}_1p_1 + \cdots + \tilde{q}_kp_k$ mit $\tilde{q}_i = a^{-1}q_i \in \mathbb{K}[z]$. \square

VI.4.6. LEMMA (Irreduzible Polynome sind prim). *Seien $q_1, q_2 \in \mathbb{K}[z]$ beliebige Polynome und $p \in \mathbb{K}[z]$ irreduzibel, sodass $p|q_1q_2$. Dann gilt $p|q_1$ oder $p|q_2$.*

BEWEIS. Wir nehmen an, dass p nicht q_1 teilt und werden $p|q_2$ zeigen. Beachte, dass p und q_1 teilerfremd sind, denn wegen der Irreduzibilität von p sind alle nicht konstanten Teiler von p von der Form cp mit $0 \neq c \in \mathbb{K}$, diese können aber nicht q_1 teilen, da q_1 nicht von p geteilt wird. Nach Lemma VI.4.5 existieren daher $a, b \in \mathbb{K}[z]$, sodass $ap + bq_1 = 1$. Wir erhalten $apq_2 + bq_1q_2 = q_2$, also $p|q_2$, denn p teilt offensichtlich den ersten Summanden, apq_2 , aber auch den zweiten, bq_1q_2 , denn nach Voraussetzung gilt $p|q_1q_2$. \square

VI.4.7. DEFINITION (Normierte Polynome). Ein Polynom $0 \neq p \in \mathbb{K}[z]$ wird *normiert* oder *monisch* genannt, falls es führenden Koeffizient 1 hat, d.h. wenn es von der Form $p = z^n + p_{n-1}z^{n-1} + \cdots + p_1z + p_0$ ist, $p_i \in \mathbb{K}$.

VI.4.8. PROPOSITION (Eindeutige Primfaktorzerlegung). *Sei $p \in \mathbb{K}[z]$ ein Polynom mit $\deg(p) \geq 1$. Dann existieren paarweise verschiedene irreduzible normierte Polynome $q_1, \dots, q_k \in \mathbb{K}[z]$, $m_1, \dots, m_k \in \mathbb{N}$ und $0 \neq c \in \mathbb{K}$, sodass*

$$p = cq_1^{m_1} \cdots q_k^{m_k}$$

Diese Darstellung ist bis auf die Reihenfolge der Faktoren eindeutig.

BEWEIS. Ist p irreduzibel und bezeichnet $0 \neq c \in \mathbb{K}$ den führenden Koeffizienten von p , dann ist $q = c^{-1}p$ ein normiertes irreduzibles Polynom und $p = cq$ die gesuchte Darstellung. Ist p nicht irreduzibel, dann existieren $p_1, p_2 \in \mathbb{K}[z]$ mit $\deg(p_i) \geq 1$, sodass $p = p_1p_2$ und daher auch $\deg(p_i) < \deg(p)$. Mittels Induktion nach $\deg(p)$ dürfen wir daher annehmen, dass normierte irreduzible Polynome q_i und $0 \neq c_i \in \mathbb{K}$ existieren, sodass $p_1 = c_1q_1 \cdots q_l$ und $p_2 = c_2q_{l+1} \cdots q_n$. Fassen wir in $p = p_1p_2 = c_1c_2q_1 \cdots q_n$ gleiche Faktoren zu Potenzen zusammen, so erhalten wir die gewünschte Darstellung von p mit $c = c_1c_2$. Damit ist die Existenz der Primfaktorzerlegung bewiesen.

Um auch die Eindeutigkeit der Darstellung zu zeigen, seien nun q_i, \tilde{q}_i normierte irreduzible Polynome und $0 \neq c \in \mathbb{K}$, $0 \neq \tilde{c} \in \mathbb{K}$, sodass

$$cq_1 \cdots q_l = \tilde{c}\tilde{q}_1 \cdots \tilde{q}_n.$$

Aus der Normiertheit der q_i und \tilde{q}_i folgt $c = \tilde{c}$ und daher

$$q_1 \cdots q_l = \tilde{q}_1 \cdots \tilde{q}_n.$$

Nach Lemma VI.4.6 teilt q_1 wenigstens eines der Polynome \tilde{q}_i . O.B.d.A. $q_1|\tilde{q}_1$. Da q_1 und \tilde{q}_1 beide normiert und irreduzibel sind, folgt $q_1 = \tilde{q}_1$. Mit Hilfe der schon früher verwendeten Kürzungsregel schließen wir $q_2 \cdots q_l = \tilde{q}_2 \cdots \tilde{q}_n$. Induktiv fortfahrend erhalten wir $l = n$ und, nach Umm Nummerieren der \tilde{q}_i , auch $q_i = \tilde{q}_i$. \square

VI.4.9. BEISPIEL (Irreduzible Polynome über alg. abg. \mathbb{K}). Ist \mathbb{K} algebraisch abgeschlossen, dann zerfällt jedes nicht konstante Polynom in Linearfaktoren, also sind die irreduzible Polynome über \mathbb{K} genau die Polynome mit Grad 1. Die eindeutige Primfaktorzerlegung eines Polynoms $p \in \mathbb{K}[z]$ mit $\deg(p) \geq 1$ ist daher von der Form

$$p = c(z - \lambda_1)^{m_1} \cdots (z - \lambda_k)^{m_k},$$

wobei $\lambda_1, \dots, \lambda_k$ die Nullstellen, m_1, \dots, m_k ihre algebraischen Vielfachheiten und $0 \neq c \in \mathbb{K}$ den führenden Koeffizienten von p bezeichnen.

VI.4.10. BEISPIEL (Irreduzible Polynome über \mathbb{R}). Neben den Polynomen ersten Grades, siehe Beispiel VI.4.3,

$$q = c(z - \lambda), \quad c, \lambda \in \mathbb{K}, \quad c \neq 0, \quad (\text{VI.33})$$

gibt es noch einen zweiten Typ irreduzibler reeller Polynome, nämlich

$$q = c((z - \alpha)^2 + \beta^2), \quad c, \lambda, \alpha, \beta \in \mathbb{R}, \quad \beta > 0, \quad c \neq 0. \quad (\text{VI.34})$$

Dieses Polynom q besitzt zwei konjugiert komplexe Nullstellen, $\lambda = \alpha \pm \mathbf{i}\beta$, hat keine reelle Nullstelle, und ist daher irreduzibel. Jedes irreduzible reelle Polynom, p , ist von der Form (VI.33) oder (VI.34). Hat p eine reelle Nullstelle, dann muss es von der Form (VI.33) sein. Hat p keine reelle Nullstelle, dann existiert nach dem Fundamentalsatz der Algebra eine komplexe Nullstelle, $p(\mu) = 0$, wobei $\mu \in \mathbb{C} \setminus \mathbb{R}$. Da p reelle Koeffizienten hat, folgt $p(\bar{\mu}) = \overline{p(\mu)} = 0$, also ist $\bar{\mu}$ eine weitere Nullstelle von p , die verschieden von μ ist. Es existiert daher ein Polynom $c \in \mathbb{C}[z]$, sodass $p = c(z - \mu)(z - \bar{\mu})$. Schreiben wir $\mu = \alpha + \mathbf{i}\beta$, dann gilt $(z - \mu)(z - \bar{\mu}) = (z - \alpha)^2 + \beta^2$. Da dies ein reelles Polynom ist, folgt $c \in \mathbb{R}[z]$, und wegen der Irreduzibilität von p muss c daher konstant sein. Somit hat p die Gestalt (VI.34). Die Primfaktorzerlegung eines beliebigen reellen Polynoms p mit $\deg(p) \geq 1$ ist daher von der Form

$$p = c(z - \lambda_1)^{m_1} \cdots (z - \lambda_k)^{m_k} ((z - \alpha_1)^2 + \beta_1^2)^{\tilde{m}_1} \cdots ((z - \alpha_l)^2 + \beta_l^2)^{\tilde{m}_l}$$

wobei $c, \lambda_i, \alpha_i, \beta_i \in \mathbb{R}$, $\beta_i > 0$ und $c \neq 0$. Dabei sind $\lambda_1, \dots, \lambda_k$ die reellen Nullstellen, m_1, \dots, m_k ihre algebraischen Vielfachheiten, $\alpha_1 + \mathbf{i}\beta_1, \dots, \alpha_l + \mathbf{i}\beta_l$ die komplexen Nullstellen mit positivem Imaginärteil und $\tilde{m}_1, \dots, \tilde{m}_l$ ihre algebraischen Vielfachheiten. Fassen wir p als komplexes Polynom auf, dann gilt

$$p = c(z - \lambda_1)^{m_1} \cdots (z - \lambda_k)^{m_k} (z - (\alpha_1 + \mathbf{i}\beta_1))^{\tilde{m}_1} (z - (\alpha_1 - \mathbf{i}\beta_1))^{\tilde{m}_1} \cdots (z - (\alpha_l + \mathbf{i}\beta_l))^{\tilde{m}_l} (z - (\alpha_l - \mathbf{i}\beta_l))^{\tilde{m}_l}$$

d.h. die komplexen Nullstellen $\alpha_1 - \mathbf{i}\beta_1, \dots, \alpha_l - \mathbf{i}\beta_l$ haben die gleichen algebraischen Vielfachheiten, $\tilde{m}_1, \dots, \tilde{m}_k$, wie die dazu komplex konjugierten Nullstellen.

VI.4.11. SATZ (Primärzerlegung). Sei V ein endlich dimensionaler \mathbb{K} -Vektorraum, $\varphi: V \rightarrow V$ linear und $p \in \mathbb{K}[z]$ ein Polynom mit $\deg(p) \geq 1$, sodass

$p(\varphi) = 0$. Bezeichnet $p = cp_1^{m_1} \cdots p_k^{m_k}$ die eindeutige Primfaktorzerlegung von p , dann zerfällt V in eine invariante direkte Summe:

$$V = \ker(p_1^{m_1}(\varphi)) \oplus \cdots \oplus \ker(p_k^{m_k}(\varphi)), \quad \varphi(\ker(p_i^{m_i}(\varphi))) \subseteq \ker(p_i^{m_i}(\varphi)).$$

BEWEIS. Bezeichne $\hat{p}_i := p/p_i^{m_i} = p_1^{m_1} \cdots p_{i-1}^{m_{i-1}} p_{i+1}^{m_{i+1}} \cdots p_k^{m_k} \in \mathbb{K}[z]$. Da die Polynome $\hat{p}_1, \dots, \hat{p}_k$ teilerfremd sind, existieren Polynome $q_i \in \mathbb{K}[z]$, sodass $1 = q_1 \hat{p}_1 + \cdots + q_k \hat{p}_k$. Setzen wir $\pi_i := (q_i \hat{p}_i)(\varphi): V \rightarrow V$, $i = 1, \dots, k$, so erhalten wir

$$\text{id}_V = \pi_1 + \cdots + \pi_k. \quad (\text{VI.35})$$

Für $i \neq j$ ist $a := p/(p_i^{m_i} p_j^{m_j}) \in \mathbb{K}[z]$ und $ap = p^2/(p_i^{m_i} p_j^{m_j}) = \hat{p}_i \hat{p}_j$, also $aq_i q_j p = q_i \hat{p}_i q_j \hat{p}_j$ und daher $\pi_i \circ \pi_j = (q_i \hat{p}_i)(\varphi) \circ (q_j \hat{p}_j)(\varphi) = (q_i \hat{p}_i q_j \hat{p}_j)(\varphi) = (aq_i q_j p)(\varphi) = (aq_i q_j)(\varphi) \circ p(\varphi) = 0$. Zusammen mit (VI.35) folgt

$$\pi_i^2 = \pi_i \quad \text{und} \quad \pi_i \circ \pi_j = 0, \quad i \neq j.$$

Nach Proposition VI.1.16 erhalten wir die Zerlegung

$$V = \text{img}(\pi_1) \oplus \cdots \oplus \text{img}(\pi_k), \quad \varphi(\text{img}(\pi_i)) \subseteq \text{img}(\pi_i),$$

wobei die Invarianz aus $\varphi \circ \pi_i = (zq_i \hat{p}_i)(\varphi) = (q_i \hat{p}_i z)(\varphi) = \pi_i \circ \varphi$ folgt. Es bleibt

$$\text{img}(\pi_i) = \ker(p_i^{m_i}(\varphi))$$

zu zeigen. Da $p_i^{m_i} q_i \hat{p}_i = q_i p$ folgt $p_i^{m_i}(\varphi) \circ \pi_i = (p_i^{m_i} q_i \hat{p}_i)(\varphi) = (q_i p)(\varphi) = q_i(\varphi) \circ p(\varphi) = 0$ und daher die Inklusion $\text{img}(\pi_i) \subseteq \ker(p_i^{m_i}(\varphi))$. Da die Polynome $p_i^{m_i}$ und $q_i \hat{p}_i$ teilerfremd sind, existieren Polynome $b, c \in \mathbb{K}[z]$, sodass $1 = bp_i^{m_i} + cq_i \hat{p}_i$, also $\text{id}_V = b(\varphi) \circ p_i^{m_i}(\varphi) + \pi_i \circ c(\varphi)$, woraus wir auch die umgekehrte Inklusion, $\ker(p_i^{m_i}(\varphi)) \subseteq \text{img}(\pi_i)$, erhalten. \square

VI.4.12. BEMERKUNG. Ist $\varphi: V \rightarrow V$ linear und bezeichnen $\lambda_1, \dots, \lambda_k \in \mathbb{K}$ alle verschiedenen Eigenwerte von φ , dann hat die eindeutige Primfaktorzerlegung des charakteristischen Polynoms von φ die Form

$$p = c(z - \lambda_1)^{m_1} \cdots (z - \lambda_k)^{m_k} p_1^{\tilde{m}_1} \cdots p_l^{\tilde{m}_l},$$

wobei m_1, \dots, m_k die algebraischen Vielfachheiten der Eigenwerte bezeichnen. Die Zerlegung aus Satz VI.4.11,

$$V = \underbrace{\ker((\varphi - \lambda_1 \text{id}_V)^{m_1})}_{\tilde{E}_{\lambda_1}} \oplus \cdots \oplus \underbrace{\ker((\varphi - \lambda_k \text{id}_V)^{m_k})}_{\tilde{E}_{\lambda_k}} \oplus \underbrace{\ker(p_1^{\tilde{m}_1}(\varphi)) \oplus \cdots \oplus \ker(p_l^{\tilde{m}_l}(\varphi))}_W$$

verfeinert die Zerlegung aus Satz VI.3.7.

VI.4.13. BEISPIEL. Die reelle Matrix

$$A = \begin{pmatrix} 0 & 1 & 7 & 0 \\ -1 & 0 & 2 & 4 \\ 0 & 0 & 0 & 2 \\ 0 & 0 & -2 & 0 \end{pmatrix}$$

hat charakteristisches Polynom $p = (z^2 + 1)(z^2 + 4)$ und daher keine reellen Eigenwerte. Es gilt:

$$\begin{aligned} \ker(A^2 + I_4) &= \ker \begin{pmatrix} 0 & 0 & 2 & 18 \\ 0 & 0 & -15 & 4 \\ 0 & 0 & -3 & 0 \\ 0 & 0 & 0 & -3 \end{pmatrix} = \left\langle \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} \right\rangle \\ \ker(A^2 + 4I_4) &= \ker \begin{pmatrix} 3 & 0 & 2 & 18 \\ 0 & 3 & -15 & 4 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} = \left\langle \begin{pmatrix} -2 \\ 15 \\ 3 \\ 0 \end{pmatrix}, \begin{pmatrix} -18 \\ -4 \\ 0 \\ 3 \end{pmatrix} \right\rangle \end{aligned}$$

Somit

$$S^{-1}AS = \left(\begin{array}{cc|cc} 0 & 1 & & \\ -1 & 0 & & \\ \hline & & 0 & 2 \\ & & -2 & 0 \end{array} \right), \quad S = \begin{pmatrix} 1 & 0 & -2 & -18 \\ 0 & 1 & 15 & -4 \\ 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 3 \end{pmatrix}.$$

VI.4.14. DEFINITION. Für $\alpha, \beta \in \mathbb{R}$ mit $\beta > 0$ definieren wir den reellen Jordanblock zum komplexen Eigenwert $\alpha + \mathbf{i}\beta$ durch

$$J_{2m}(\alpha, \beta) := \begin{pmatrix} \alpha & -\beta & & & & \\ & \alpha & & & & \\ & & I_2 & & & \\ & & & \alpha & -\beta & \dots \\ & & & \beta & \alpha & \dots \\ & & & & & \dots \\ & & & & & & I_2 \\ & & & & & & & \alpha & -\beta \\ & & & & & & & \beta & \alpha \end{pmatrix} \in M_{2m \times 2m}(\mathbb{R})$$

Das charakteristische Polynom eines solchen Jordanblocks ist

$$p = ((z - \alpha)^2 + \beta^2)^m = (z - (\alpha + \mathbf{i}\beta))^m (z - (\alpha - \mathbf{i}\beta))^m$$

und hat daher zwei komplex konjugierte Nullstellen, $\lambda = \alpha \pm \mathbf{i}\beta$, beide mit algebraischer Vielfachheit m .

VI.4.15. LEMMA. Sei V ein endlich dimensionaler reeller Vektorraum und $\varphi: V \rightarrow V$ linear mit charakteristischem Polynom

$$p = ((z - \alpha)^2 + \beta^2)^m,$$

wobei $\alpha, \beta \in \mathbb{R}$ und $\beta > 0$. Dann existiert eine Basis B von V , sodass die Matrixdarstellung $[\varphi]_{BB}$ folgende Blockdiagonalgestalt hat:

$$[\varphi]_{BB} = \begin{pmatrix} J_{2m_1}(\alpha, \beta) & & \\ & \dots & \\ & & J_{2m_n}(\alpha, \beta) \end{pmatrix}$$

Dabei sind $m_1, \dots, m_n \in \mathbb{N}$ und $m_1 + \dots + m_n = m$. Für die Anzahl der auftretenden Jordanblöcke gilt weiters:

$$2\#\{j \mid m_j \geq r\} = \dim \ker \left(((\varphi - \alpha \operatorname{id}_V)^2 + \beta^2 \operatorname{id}_V)^r \right) \\ - \dim \ker \left(((\varphi - \alpha \operatorname{id}_V)^2 + \beta^2 \operatorname{id}_V)^{r-1} \right)$$

BEWEIS. O.B.d.A. sei $V = \mathbb{R}^{2m}$ und φ durch eine Matrix $A \in M_{2m \times 2m}(\mathbb{R})$ gegeben, $\varphi(x) = Ax$. Wir fassen A nun als komplexe Matrix auf, $A \in M_{2m \times 2m}(\mathbb{C})$, und betrachten die damit assoziierte komplex lineare Abbildung $\varphi_{\mathbb{C}}: \mathbb{C}^{2m} \rightarrow \mathbb{C}^{2m}$, $\varphi_{\mathbb{C}}(x) := Ax$. Für ihr charakteristisches Polynom gilt

$$p = (z - \lambda)^m (z - \bar{\lambda})^m.$$

wobei $\lambda := \alpha + \mathbf{i}\beta$ und $\bar{\lambda} = \alpha - \mathbf{i}\beta$. Nach Satz VI.3.7 gilt

$$\mathbb{C}^{2m} = \tilde{E}_{\lambda} \oplus \tilde{E}_{\bar{\lambda}}. \quad (\text{VI.36})$$

Da A reelle Eintragungen hat gilt

$$\overline{\tilde{E}_{\lambda}} = \tilde{E}_{\bar{\lambda}} \quad \text{und} \quad \overline{\tilde{E}_{\bar{\lambda}}} = \tilde{E}_{\lambda}, \quad (\text{VI.37})$$

denn für $x \in \tilde{E}_{\bar{\lambda}}$ ist $(A - \lambda I_{2m})^m x = 0$, also

$$(A - \bar{\lambda} I_{2m})^m \bar{x} = (\bar{A} - \bar{\lambda} I_{2m})^m \bar{x} = \overline{(A - \lambda I_{2m})^m x} = \bar{0} = 0,$$

und daher $\bar{x} \in \tilde{E}_{\lambda}$. Nach Satz VI.3.19 existieren eine Basis c_1, \dots, c_m von \tilde{E}_{λ} und $\varepsilon_j \in \{0, 1\}$, sodass

$$Ac_j = \lambda c_j + \varepsilon_{j-1} c_{j-1}, \quad j = 1, \dots, m,$$

wobei $\varepsilon_0 := 0$. Da $\bar{A} = A$ und $\bar{\varepsilon}_j = \varepsilon_j$, folgt daraus

$$A\bar{c}_j = \bar{\lambda} \bar{c}_j + \varepsilon_{j-1} \bar{c}_{j-1}, \quad j = 1, \dots, m.$$

Setzen wir

$$b_j := \frac{1}{2}(c_j + \bar{c}_j), \quad b'_j := \frac{1}{2\mathbf{i}}(c_j - \bar{c}_j), \quad j = 1, \dots, m,$$

so gilt

$$\bar{b}_j = b_j, \quad \bar{b}'_j = b'_j, \quad c_j = b_j + \mathbf{i}b'_j, \quad \bar{c}_j = b_j - \mathbf{i}b'_j,$$

sowie

$$\lambda c_j = (\alpha b_j - \beta b'_j) + \mathbf{i}(\beta b_j + \alpha b'_j).$$

Durch einfache Rechnung folgt

$$Ab_j = \alpha b_j + \beta b'_j + \varepsilon_{j-1} b_{j-1}, \quad Ab'_j = -\beta b_j + \alpha b'_j + \varepsilon_{j-1} b'_{j-1}. \quad (\text{VI.38})$$

Wegen (VI.37) bildet $\bar{c}_1, \dots, \bar{c}_m$ eine Basis von $\tilde{E}_{\bar{\lambda}}$, also ist $c_1, \dots, c_m, \bar{c}_1, \dots, \bar{c}_m$ eine Basis von \mathbb{C}^{2m} , siehe (VI.36). Da $b_j + \mathbf{i}b'_j = c_j$ und $b_j - \mathbf{i}b'_j = \bar{c}_j$ ist auch

$$B = b_1, b'_1, b_2, b'_2, \dots, b_m, b'_m$$

eine Basis von \mathbb{C}^{2m} . Nach (VI.38) gilt

$$[\varphi_{\mathbb{C}}]_{BB} = \begin{pmatrix} \alpha & -\beta & & & & \\ \beta & \alpha & \varepsilon_1 I_2 & & & \\ & & \alpha & -\beta & \ddots & \\ & & \beta & \alpha & & \\ & & & & \ddots & \varepsilon_{m-1} I_2 \\ & & & & & \alpha & -\beta \\ & & & & & \beta & \alpha \end{pmatrix}$$

Da $\bar{b}_j = b_j$ und $\bar{b}'_j = b'_j$ liegen alle Basisvektoren von B in $\mathbb{R}^{2m} \subseteq \mathbb{C}^{2m}$, also bildet B auch eine Basis des reellen Vektorraums \mathbb{R}^{2m} und $[\varphi]_{BB} = [\varphi_{\mathbb{C}}]_{BB}$ hat die gewünschte Gestalt. Die Formel für die Anzahl der verschiedenen Jordanblöcke folgt aus, siehe Aufgabe 72,

$$\dim \ker \left(((\varphi - \alpha \operatorname{id}_V)^2 + \beta^2 \operatorname{id}_V)^r \right) = 2\#\{j \mid m_j \geq 1\} + \cdots + 2\#\{j \mid m_j \geq r\}. \quad \square$$

VI.4.16. SATZ (Reelle Jordan'sche Normalform). *Sei V ein endlich dimensionaler reeller Vektorraum, $\varphi: V \rightarrow V$ linear und*

$$p = (z - \lambda_1)^{m_1} \cdots (z - \lambda_k)^{m_k} ((z - \alpha_1)^2 + \beta_1^2)^{\tilde{m}_1} \cdots ((z - \alpha_l)^2 + \beta_l^2)^{\tilde{m}_l}$$

die eindeutige Primfaktorzerlegung des charakteristischen Polynoms von φ , wobei $\lambda_i, \alpha_i, \beta_i \in \mathbb{R}$ und $\beta_i > 0$. Dann existiert eine Basis B von V , sodass $[\varphi]_{BB}$ Blockdiagonalgestalt hat, wobei entlang der Diagonale folgende Jordanblöcke stehen:

$$J_{m_{1,1}}(\lambda_1), \dots, J_{m_{1,n_1}}(\lambda_1), \dots, J_{m_{k,1}}(\lambda_k), \dots, J_{m_{k,n_k}}(\lambda_k), \\ J_{2\tilde{m}_{1,1}}(\alpha_1, \beta_1), \dots, J_{2\tilde{m}_{1,\tilde{n}_1}}(\alpha_1, \beta_1), \dots, J_{2\tilde{m}_{l,1}}(\alpha_l, \beta_l), \dots, J_{2\tilde{m}_{l,\tilde{n}_l}}(\alpha_l, \beta_l).$$

Dabei sind $m_{i,j}, \tilde{m}_{i,j}, n_i, \tilde{n}_i \in \mathbb{N}$ und es gilt

$$m_{i,1} + \cdots + m_{i,n_i} = m_i \quad \text{sowie} \quad \tilde{m}_{i,1} + \cdots + \tilde{m}_{i,\tilde{n}_i} = \tilde{m}_i.$$

Für die Anzahl der auftretenden Jordanblöcke gilt weiters:

$$\#\{j \mid m_{i,j} \geq r\} = \dim \ker((\varphi - \lambda_i \operatorname{id}_V)^r) - \dim \ker((\varphi - \lambda_i \operatorname{id}_V)^{r-1}) \\ 2\#\{j \mid \tilde{m}_{i,j} \geq r\} = \dim \ker(((\varphi - \alpha_i \operatorname{id}_V)^2 + \beta_i^2 \operatorname{id}_V)^r) \\ - \dim \ker(((\varphi - \alpha_i \operatorname{id}_V)^2 + \beta_i^2 \operatorname{id}_V)^{r-1})$$

Insbesondere ist diese Matrixdarstellung, $[\varphi]_{BB}$, bis auf die Reihenfolge der Jordanblöcke eindeutig bestimmt. Zwei lineare Abbildungen zwischen reellen Vektorräumen sind genau dann ähnlich, wenn ihre Jordan'schen Normalformen, bis auf die Reihenfolge der Jordanblöcke, übereinstimmen.

BEWEIS. Nach dem Satz von Cayley–Hamilton gilt $p(\varphi) = 0$. Aufgrund der Primärzerlegung in Satz VI.4.11 genügt es daher Abbildungen mit charakteristischem Polynom $p = (z - \lambda)^m$ bzw. $p = ((z - \alpha)^2 + \beta^2)^m$ zu betrachten. Erstere haben wir in Satz VI.3.19 behandelt, die anderen in Lemma VI.4.15. \square

VI.4.17. KOROLLAR. Sei V ein endlich dimensionaler reeller Vektorraum und $\varphi: V \rightarrow V$ linear. Dann existiert eine Basis B von V , sodass die Matrixdarstellung $[\varphi]_{BB}$ fast obere Dreiecksgestalt hat, soll heißen,

$$[\varphi]_{BB} = \begin{pmatrix} \lambda_1 & * & * & * & \cdots & \cdots & \cdots & * \\ & \ddots & \ddots & \ddots & \ddots & & & \vdots \\ & & \lambda_k & * & * & * & & \vdots \\ & & & \alpha_1 & -\beta_1 & * & \ddots & \vdots \\ & & & \beta_1 & \alpha_1 & * & \ddots & * \\ & & & & & \ddots & * & * \\ & & & & & & \alpha_l & -\beta_l \\ & & & & & & \beta_l & \alpha_l \end{pmatrix}$$

wobei $\lambda_1, \dots, \lambda_k, \alpha_l + \mathbf{i}\beta_1, \dots, \alpha_l + \mathbf{i}\beta_l$ alle komplexen Eigenwerte mit nicht negativen Imaginärteil, ihrer algebraischen Vielfachheit entsprechend oft gelistet, bezeichnen, $\lambda_i, \alpha_i, \beta_i \in \mathbb{R}$, $\beta_i > 0$. In anderen Worten:

$$p_\varphi = (z - \lambda_1) \cdots (z - \lambda_k) ((z - \alpha_1)^2 + \beta_1^2) \cdots ((z - \alpha_l)^2 + \beta_l^2).$$

VI.4.18. BEMERKUNG (Minimalpolynom). Sei $A \in M_{n \times n}(\mathbb{K})$ eine quadratische Matrix. Da $\dim(M_{n \times n}(\mathbb{K})) = n^2$, sind die Matrizen $I_n, A, A^2, A^3, \dots, A^{n^2}$ linear abhängig in $M_{n \times n}(\mathbb{K})$. Es existieren daher Skalare $q_0, \dots, q_{n^2} \in \mathbb{K}$, die nicht alle verschwinden, sodass $\sum_{i=0}^{n^2} q_i A^i = 0$. Für $q := \sum_{i=0}^{n^2} q_i z^i$ gilt daher $0 \neq q \in \mathbb{K}[z]$ und $q(A) = 0$. Somit ist

$$J_A := \{p \in \mathbb{K}[z] \mid p(A) = 0\} \neq \{0\}.$$

Offensichtlich bildet J_A einen Teilraum von $\mathbb{K}[z]$, und für jedes $p \in J_A$ und $r \in \mathbb{K}[z]$ gilt auch $rp \in J_A$, denn $(rp)(A) = r(A)p(A) = r(A)0 = 0$. Solche Teilräume werden *Ideale* genannt. Sei nun $0 \neq m_A \in J_A$ normiert mit minimalem Grad, d.h. $\deg(m_A) \leq \deg(p)$, für jedes $0 \neq p \in J_A$. Es gilt dann

$$J_A = \{sm_A \mid s \in \mathbb{K}[z]\},$$

denn zu $p \in J_A$ existieren $s, r \in \mathbb{K}[z]$ mit $p = sm_A + r$ und $\deg(r) < \deg(m_A)$, folglich $r = p - sm_A \in J_A$, und wegen der Minimalität von $\deg(m_A)$ muss $r = 0$ gelten, also $p = sm_A$. Solche Ideale, die von einem Element, m_A , erzeugt werden, heißen *Hauptideale*. Das Argument oben zeigt, dass jedes Ideal von $\mathbb{K}[z]$ ein Hauptideal ist. Insbesondere sehen wir, dass das Polynom m_A eindeutig bestimmt ist, es wird das *Minimalpolynom* von A genannt. Nach dem Satz von Caley–Hamilton liegt auch das charakteristische Polynom von A in diesem Ideal, $p_A \in J_A$, es wird daher vom Minimalpolynom geteilt, $m_A \mid p_A$.

VII. Euklidische und unitäre Vektorräume

Wir werden in diesem Abschnitt Vektorräume, die mit einem inneren Produkt (Skalarprodukt) ausgestattet sind, untersuchen. Als prototypisches Beispiel dient das bekannte standard Euklidische innere Produkt auf \mathbb{R}^n ,

$$\langle x, y \rangle = x^t y = x_1 y_1 + \cdots + x_n y_n, \quad x, y \in \mathbb{R}^n.$$

Damit lassen sich Längen von Vektoren, Winkel zwischen Vektoren, orthogonale Komplemente, Isometrien (d.h. Längen und Winkel bewahrende lineare Abbildungen), u.v.a.m. definieren und studieren. Folgende Eigenschaften dieses inneren Produkts haben sich als wesentlich herausgestellt:

- (a) *Bilinearität*: Die Abbildung $\langle -, - \rangle: \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$, $\langle x, y \rangle = x^t y$, ist bilinear, d.h. linear in jeder der beiden Eintragungen. Es gilt daher $\langle x + x', y \rangle = \langle x, y \rangle + \langle x', y \rangle$, $\langle x, y + y' \rangle = \langle x, y \rangle + \langle x, y' \rangle$ und $\langle \lambda x, y \rangle = \lambda \langle x, y \rangle = \langle x, \lambda y \rangle$, für alle $x, x', y, y' \in \mathbb{R}^n$ und $\lambda \in \mathbb{K}$.
- (b) *Symmetrie*: $\langle x, y \rangle = \langle y, x \rangle$, für alle $x, y \in \mathbb{R}^n$.
- (c) *Positivität*: $\langle x, x \rangle > 0$, für alle $0 \neq x \in \mathbb{R}^n$.

VII.1. Symmetrische Bilinearformen. Eine naheliegende Verallgemeinerung der Eigenschaften (a) und (b) oben führt zum Begriff der symmetrischen Bilinearform auf allgemeinen Vektorräumen.

VII.1.1. DEFINITION (Symmetrische Bilinearformen). Sei V ein Vektorraum über einem Körper \mathbb{K} . Unter einer *Bilinearform* auf V verstehen wir eine Abbildung

$$\beta: V \times V \rightarrow \mathbb{K},$$

die linear in jeder Eintragung ist, d.h.

$$\beta(v + v', w) = \beta(v, w) + \beta(v', w), \quad \beta(\lambda v, w) = \lambda \beta(v, w)$$

und

$$\beta(v, w + w') = \beta(v, w) + \beta(v, w'), \quad \beta(v, \lambda w) = \lambda \beta(v, w),$$

für alle $v, v', w, w' \in V$ und $\lambda \in \mathbb{K}$. Gilt darüber hinaus

$$\beta(v, w) = \beta(w, v),$$

so wird β eine *symmetrische* Bilinearform genannt.

VII.1.2. BEMERKUNG. Ist $\beta: V \times V \rightarrow \mathbb{K}$ linear in der ersten Eintragung und symmetrisch, dann muss es auch linear in der zweiten Eintragung sein, siehe Aufgabe 80.

Die Menge der Bilinearformen auf V bildet einen Vektorraum bezüglich der Operationen

$$(\beta_1 + \beta_2)(v, w) := \beta_1(v, w) + \beta_2(v, w), \quad (\lambda \beta)(v, w) = \lambda \beta(v, w).$$

Dabei sind β, β_1, β_2 Bilinearformen auf V , $\lambda \in \mathbb{K}$ und $v, w \in V$. Die Menge der symmetrischen Bilinearformen auf V bildet einen Teilraum und ist daher selbst ein Vektorraum über \mathbb{K} , siehe Aufgabe 79.

VII.1.3. BEISPIEL (Mit Matrizen assoziierte Bilinearformen auf \mathbb{K}^n). Jede quadratische Matrix $A \in M_{n \times n}(\mathbb{K})$ liefert eine Bilinearform auf \mathbb{K}^n ,

$$\beta_A: \mathbb{K}^n \times \mathbb{K}^n \rightarrow \mathbb{K}, \quad \beta_A(x, y) := x^t A y,$$

denn $(x_1 + x_2)^t A y = x_1^t A y + x_2^t A y$, $x^t A (y_1 + y_2) = x^t A y_1 + x^t A y_2$ und $(\lambda x)^t A y = \lambda(x^t A y) = x^t A (\lambda y)$. Beachte, dass die Matrix A durch die Bilinearform β_A eindeutig bestimmt ist, denn $A_{ij} = \beta_A(e_i, e_j)$. Diese Bilinearform β_A ist genau dann symmetrisch, wenn die Matrix A symmetrisch ist, d.h. $A^t = A$, denn $\beta_A(y, x) = y^t A x = (y^t A x)^t = x^t A^t y = \beta_{A^t}(x, y)$. Umgekehrt lässt sich jede (symmetrische) Bilinearform $\beta: \mathbb{K}^n \times \mathbb{K}^n \rightarrow \mathbb{K}$ durch eine (symmetrische) Matrix A beschreiben, denn für alle $x, y \in \mathbb{K}^n$ gilt

$$\beta(x, y) = \beta\left(\sum_{i=1}^n x_i e_i, \sum_{j=1}^n y_j e_j\right) = \sum_{i,j=1}^n x_i y_j \beta(e_i, e_j) = \sum_{i,j=1}^n x_i y_j A_{ij} = \beta_A(x, y),$$

also $\beta = \beta_A$, wobei $A_{ij} := \beta(e_i, e_j)$. Beachte auch $\beta_{A+A'} = \beta_A + \beta_{A'}$ und $\beta_{\lambda A} = \lambda \beta_A$, d.h. die Zuordnung $A \leftrightarrow \beta_A$ liefert einen linearen Isomorphismus zwischen dem Vektorraum der (symmetrischen) $(n \times n)$ -Matrizen und dem Vektorraum der (symmetrischen) Bilinearformen auf \mathbb{K}^n . Insbesondere hat der Vektorraum der Bilinearformen auf \mathbb{K}^n Dimension n^2 und der Teilraum der symmetrischen Bilinearformen auf \mathbb{K}^n hat Dimension $n(n+1)/2$.

VII.1.4. BEISPIEL. Das Euklidische innere Produkt, $\langle -, - \rangle: \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$,

$$\langle x, y \rangle = \sum_{i=1}^n x_i y_i = x^t y = x^t \begin{pmatrix} 1 & & \\ & \ddots & \\ & & 1 \end{pmatrix} y,$$

ist eine symmetrische Bilinearform auf \mathbb{R}^n .

VII.1.5. BEISPIEL. Die Lorentz Metrik, $g: \mathbb{R}^4 \times \mathbb{R}^4 \rightarrow \mathbb{R}$,

$$g(x, y) = -x_0 y_0 + x_1 y_1 + x_2 y_2 + x_3 y_3 = x^t \begin{pmatrix} -1 & & & \\ & 1 & & \\ & & 1 & \\ & & & 1 \end{pmatrix} y,$$

ist eine symmetrische Bilinearform auf \mathbb{R}^4 .

VII.1.6. BEISPIEL. Die Abbildung

$$\beta: M_{m \times n}(\mathbb{K}) \times M_{m \times n}(\mathbb{K}) \rightarrow \mathbb{K}, \quad \beta(A, B) := \text{tr}(A^t B)$$

ist eine symmetrische Bilinearform auf dem Vektorraum der Matrizen, $M_{m \times n}(\mathbb{K})$.

VII.1.7. BEISPIEL. Sei $I \subseteq \mathbb{R}$ ein kompaktes Intervall und bezeichnen $C^0(I; \mathbb{R})$ den Vektorraum der stetigen Funktionen auf I . Dann bildet

$$\beta: C^0(I; \mathbb{R}) \times C^0(I; \mathbb{R}) \rightarrow \mathbb{R}, \quad \beta(f, g) := \int_I f(t)g(t)dt,$$

eine symmetrische Bilinearform auf $C^0(I; \mathbb{R})$.

VII.1.8. DEFINITION (Matrixdarstellung). Sei $\beta: V \times V \rightarrow \mathbb{K}$ eine Bilinearform auf einem n -dimensionalen \mathbb{K} -Vektorraum V . Ist $B = (b_1, \dots, b_n)$ eine geordnete Basis von V , dann wird die Matrix $[\beta]_B \in M_{n \times n}(\mathbb{K})$,

$$([\beta]_B)_{ij} := \beta(b_i, b_j), \quad 1 \leq i, j \leq n,$$

als *Matrixdarstellung von β bezüglich der Basis B* bezeichnet.

VII.1.9. PROPOSITION. Sei V ein n -dimensionaler \mathbb{K} -Vektorraum. Jede geordnete Basis B von V liefert einen linearen Isomorphismus, $\beta \leftrightarrow [\beta]_B$, zwischen dem Vektorraum der (symmetrischen) Bilinearformen auf V und dem Vektorraum der (symmetrischen) $(n \times n)$ -Matrizen. Insbesondere ist β durch die Matrix $[\beta]_B$ eindeutig bestimmt,

$$\beta(v, w) = [v]_B^t [\beta]_B [w]_B, \quad v, w \in V,$$

und zu jeder (symmetrischen) $(n \times n)$ -Matrix A existiert eine eindeutig bestimmte (symmetrische) Bilinearform β auf V , sodass $[\beta]_B = A$. Ist C eine weitere geordnete Basen von V , dann gilt

$$[\beta]_B = T_{CB}^t [\beta]_C T_{CB},$$

wobei $T_{CB} \in \text{GL}_n(\mathbb{K})$ die Matrix zum Basiswechsel von B nach C bezeichnet.

BEWEIS. Die Zuordnung $\beta \mapsto [\beta]_B$ ist linear, d.h. für Bilinearformen β, β_1, β_2 auf V und $\lambda \in \mathbb{K}$ gilt

$$[\beta_1 + \beta_2]_B = [\beta_1]_B + [\beta_2]_B \quad \text{und} \quad [\lambda\beta]_B = \lambda[\beta]_B,$$

denn $([\beta_1 + \beta_2]_B)_{ij} = (\beta_1 + \beta_2)(b_i, b_j) = \beta_1(b_i, b_j) + \beta_2(b_i, b_j) = ([\beta_1]_B)_{ij} + ([\beta_2]_B)_{ij} = ([\beta_1]_B + [\beta_2]_B)_{ij}$, und analog $([\lambda\beta]_B)_{ij} = (\lambda\beta)(b_i, b_j) = \lambda\beta(b_i, b_j) = \lambda([\beta]_B)_{ij}$.

Für $v = \sum_{i=1}^n ([v]_B)_i b_i$ und $w = \sum_{j=1}^n ([w]_B)_j b_j$ gilt weiters

$$\begin{aligned} \beta(v, w) &= \beta\left(\sum_{i=1}^n ([v]_B)_i b_i, \sum_{j=1}^n ([w]_B)_j b_j\right) = \sum_{i,j=1}^n ([v]_B)_i ([w]_B)_j \beta(b_i, b_j) \\ &= \sum_{i,j=1}^n ([v]_B)_i ([w]_B)_j ([\beta]_B)_{ij} = [v]_B^t [\beta]_B [w]_B, \end{aligned}$$

also ist die Zuordnung $\beta \mapsto [\beta]_B$ injektiv. Um auch die Surjektivität einzusehen, sei nun $A \in M_{n \times n}(\mathbb{K})$. Definieren wir $\beta: V \times V \rightarrow \mathbb{K}$ durch $\beta(v, w) := [v]_B^t A [w]_B$, so ist β offensichtlich bilinear und $[\beta]_B = A$. Dies zeigt, dass die Zuordnung $\beta \mapsto [\beta]_B$ auch surjektiv, also ein Isomorphismus ist. Offensichtlich ist die Matrix $[\beta]_B$ genau dann symmetrisch, wenn die Bilinearform β symmetrisch ist. Für jede

weitere geordnete Basis C haben wir $[v]_C = T_{CB}[v]_B$ und $[w]_C = T_{CB}[w]_B$, somit

$$\begin{aligned} [v]_B^t [\beta]_B [w]_B &= \beta(v, w) = [v]_C^t [\beta]_C [w]_C \\ &= (T_{CB}[v]_B)^t [\beta]_C (T_{CB}[w]_B) = [v]_B^t \left(T_{CB}^t [\beta]_C T_{CB} \right) [w]_B, \end{aligned}$$

also $[\beta]_B = T_{CB}^t [\beta]_C T_{CB}$. \square

VII.1.10. BEMERKUNG. Beachte, dass sich die Matrix einer Bilinearform bei Basiswechsel anders transformiert, als die Matrix einer linearen Abbildung. Sind B, C zwei Basen von V und ist $\varphi: V \rightarrow V$ linear, dann gilt $[\varphi]_{BB} = T_{CB}^{-1} [\varphi]_{CC} T_{CB}$, wohingegen $[\beta]_B = T_{CB}^t [\beta]_C T_{CB}$, für jede Bilinearform β auf V .

VII.1.11. BEMERKUNG (Orthogonale Gruppe). Ist $\beta: V \times V \rightarrow \mathbb{K}$ eine symmetrische Bilinearform, dann bildet

$$O(V, \beta) := \{ \varphi \in \text{GL}(V) \mid \forall v, w \in V : \beta(\varphi(v), \varphi(w)) = \beta(v, w) \}$$

eine Gruppe. Ist V endlich dimensional, B eine Basis von V und $A := [\beta]_B$, dann schränkt sich der Gruppenisomorphismus $\text{GL}(V) \cong \text{GL}_n(\mathbb{K})$, $\varphi \leftrightarrow [\varphi]_{BB}$, siehe Bemerkung IV.6.17, zu einem Gruppenisomorphismus,

$$O(V, \beta) \cong \{ X \in \text{GL}_n(\mathbb{K}) \mid X^t A X = A \},$$

ein. Dies folgt aus Proposition VII.1.9, denn

$$\begin{aligned} \beta(\varphi(v), \varphi(w)) &= [\varphi(v)]_B^t [\beta]_B [\varphi(w)]_B \\ &= ([\varphi]_{BB}[v]_B)^t [\beta]_B ([\varphi]_{BB}[w]_B) = [v]_B^t \left([\varphi]_{BB}^t [\beta]_B [\varphi]_{BB} \right) [w]_B \end{aligned}$$

und $\beta(v, w) = [v]_B^t [\beta]_B [w]$.

Jede Bilinearform $\beta: V \times V \rightarrow \mathbb{K}$ definiert eine lineare Abbildung

$$\check{\beta}: V \rightarrow V^*, \quad \check{\beta}(v)(w) := \beta(v, w),$$

und jede linear Abbildung $V \rightarrow V^*$ ist von dieser Form für eine eindeutig bestimmte Bilinearform β . Die Bilinearform β ist genau dann symmetrisch, wenn die Komposition

$$V \xrightarrow{\iota} V^{**} \xrightarrow{\check{\beta}^t} V^*$$

mit $\check{\beta}$ übereinstimmt, denn $(\check{\beta}^t(\iota(v)))(w) = \iota(v)(\check{\beta}(w)) = \check{\beta}(w)(v)$. Dabei bezeichnet ι die kanonische Inklusion aus Proposition III.4.13 und $\check{\beta}^t$ die zu $\check{\beta}$ duale Abbildung. Ist V endlich dimensional, B eine Basis von V und B^* die duale Basis von V^* , dann gilt

$$[\beta]_B = [\check{\beta}]_{B^*B}^t, \tag{VII.1}$$

denn $\check{\beta}(b_i) = \sum_{j=1}^n \check{\beta}(b_i)(b_j) b_j^* = \sum_{j=1}^n \beta(b_i, b_j) b_j^* = \sum_{j=1}^n ([\beta]_B)_{ij} b_j^*$, wobei $B = (b_1, \dots, b_n)$ und $B^* = (b_1^*, \dots, b_n^*)$.

VII.1.12. DEFINITION (Rang und Kern). Sei V ein endlich dimensionaler \mathbb{K} -Vektorraum und $\beta: V \times V \rightarrow \mathbb{K}$ eine symmetrische Bilinearform. Der Teilraum

$$\ker(\beta) := \{v \in V \mid \forall w \in V : \beta(v, w) = 0\} = \ker(\check{\beta}: V \rightarrow V^*)$$

wird *Kern* oder *Radikal* der symmetrischen Bilinearform genannt. Unter dem *Rang* von β verstehen wir den Rang der linearen Abbildung $\check{\beta}: V \rightarrow V^*$, d.h.

$$\text{rank}(\beta) := \text{rank}(\check{\beta}: V \rightarrow V^*) = \text{rank}([\beta]_B),$$

wobei B eine beliebige Basis von V bezeichnet, siehe (VII.1). Beachte

$$\dim \ker(\beta) + \text{rank}(\beta) = \dim(V), \quad (\text{VII.2})$$

denn $\dim \ker(\check{\beta}) + \dim \text{img}(\check{\beta}) = \dim(V)$ nach Korollar IV.2.9.

VII.1.13. PROPOSITION. Sei V ein endlich dimensionaler Vektorraum über einem Körper \mathbb{K} und $\beta: V \times V \rightarrow \mathbb{K}$ eine symmetrische Bilinearform. Dann sind äquivalent:

- (a) Zu jedem $0 \neq v \in V$ existiert $w \in V$, sodass $\beta(v, w) \neq 0$.
- (b) Die lineare Abbildung $\check{\beta}: V \rightarrow V^*$, $\check{\beta}(v) = \beta(v, -)$, ist ein Isomorphismus.
- (c) $\ker(\beta) = \{0\}$.
- (d) $\text{rank}(\beta) = \dim(V)$.
- (e) Für eine (und dann jede) Basis B von V ist die Matrix $[\beta]_B$ invertierbar.

BEWEIS. Die Äquivalenz (a) \Leftrightarrow (c) ist offensichtlich. Da $\dim(V^*) = \dim(V)$ erhalten wir (b) \Leftrightarrow (c) \Leftrightarrow (d) aus Korollar IV.2.11. Die Äquivalenz (b) \Leftrightarrow (e) folgt aus (VII.1). \square

VII.1.14. DEFINITION (Nicht degenerierte Bilinearformen). Eine Bilinearform auf einem endlich dimensionalen Vektorraum wird *nicht-degeneriert* genannt, wenn sie die äquivalenten Eigenschaften in Proposition VII.1.13 hat. Eine symmetrische Matrix $A \in M_{n \times n}(\mathbb{K})$ wird nicht-degeneriert genannt, falls die damit assoziierte symmetrische Bilinearform $\beta_A: \mathbb{K}^n \times \mathbb{K}^n \rightarrow \mathbb{K}$, $\beta_A(x, y) = x^t A y$, nicht-degeneriert ist. Nach Proposition VII.1.13 ist dies genau dann der Fall, wenn A invertierbar ist.

VII.1.15. BEISPIEL. Das Euklidische innere Produkt aus Beispiel VII.1.4 ist nicht-degeneriert. Auch die Lorentz Metrik aus Beispiel VII.1.5 ist nicht-degeneriert. Ebenso ist die Bilinearform aus Beispiel VII.1.6 nicht-degeneriert. Für die Bilinearform aus Beispiel VII.1.7 gilt $\ker(\beta) = \{0\}$, aber $\check{\beta}: C^0(I, \mathbb{R}) \rightarrow C^0(I, \mathbb{R})^*$ ist *nicht* surjektiv. Für jedes $x \in I$ ist nämlich $\text{ev}_x: C^0(I, \mathbb{R}) \rightarrow \mathbb{R}$, $\text{ev}_x(g) := g(x)$, ein lineares Funktional, aber es existiert *kein* $f \in C^0(I, \mathbb{R})$, sodass $\text{ev}_x(g) = \int_I f(t)g(t)dt$, für alle $g \in C^0(I, \mathbb{R})$.

VII.1.16. BEMERKUNG. Ist $\beta: V \times V \rightarrow \mathbb{K}$ eine nicht-degenerierte Bilinearform auf einem endlich dimensionalen Vektorraum, V , dann gilt

$$O(V, \beta) = \{\varphi \in \text{end}(V) \mid \forall v, w \in V : \beta(\varphi(v), \varphi(w)) = \beta(v, w)\},$$

d.h. eine lineare Abbildung, die β bewahrt, ist automatisch invertierbar. Aus $\beta(\varphi(v), \varphi(w)) = \beta(v, w)$ folgt nämlich $\ker(\varphi) \subseteq \ker(\beta)$, also $\ker(\varphi) = \{0\}$ und daher $\varphi \in \text{GL}(V)$.

Ist $\beta: V \times V \rightarrow \mathbb{K}$ eine (symmetrische) Bilinearform und $W \subseteq V$ ein Teilraum, dann ist offensichtlich auch die Einschränkung

$$\beta|_W: W \times W \rightarrow \mathbb{K}, \quad \beta|_W(w, w') := \beta(w, w'), \quad w, w' \in W,$$

eine (symmetrische) Bilinearform auf W .

VII.1.17. PROPOSITION. *Sei $\beta: V \times V \rightarrow \mathbb{K}$ eine symmetrische Bilinearform auf einem endlich dimensionalen Vektorraum V . Ist W ein zu $\ker(\beta)$ komplementärer Teilraum, dann ist $\beta|_W$ nicht-degeneriert. Es existiert daher eine Basis B von V , sodass $[\beta]_B = \begin{pmatrix} A & 0 \\ 0 & 0 \end{pmatrix}$, wobei $A \in M_{k \times k}(\mathbb{K})$ invertierbar und*

$$k = \text{rank}(\beta) = \max\{\dim(W) \mid W \text{ Teilraum von } V, \text{ s.d. } \beta|_W \text{ nicht-deg.}\}.$$

BEWEIS. Sei W ein zu $\ker(\beta)$ komplementärer Teilraum, d.h. $V = W \oplus \ker(\beta)$. Sei nun $w \in \ker(\beta|_W)$, d.h. $w \in W$ und $\beta(w, w') = 0$, für alle $w' \in W$. Andererseits haben wir auch $\beta(w, u) = 0$, für alle $u \in \ker(\beta)$. Da sich jedes $v \in V$ in der Form $v = u + w'$ schreiben lässt, folgt $\beta(w, v) = 0$, für alle $v \in V$, also $w \in \ker(\beta)$. Da $w \in W$, folgt $w \in W \cap \ker(\beta) = \{0\}$, also $w = 0$. Dies zeigt $\ker(\beta|_W) = \{0\}$, nach Proposition VII.1.13 ist $\beta|_W$ daher nicht-degeneriert. Ist $\tilde{B} = (b_1, \dots, b_k)$ eine Basis von W und b_{k+1}, \dots, b_n eine Basis von $\ker(\beta)$, dann bildet $B = (b_1, \dots, b_n)$ eine Basis von V und $[\beta]_B = \begin{pmatrix} A & 0 \\ 0 & 0 \end{pmatrix}$, wobei $A \in M_{k \times k}(\mathbb{K})$. Weiters ist $A = [\beta|_W]_{\tilde{B}}$ invertierbar, denn $\beta|_W$ ist nicht-degeneriert, siehe Proposition VII.1.13. Daraus, oder via (VII.2), folgt nun auch $k = \text{rank}(\beta)$ und

$$\text{rank}(\beta) \leq \max\{\dim(W) \mid W \text{ Teilraum von } V, \text{ s.d. } \beta|_W \text{ nicht-deg.}\}. \quad (\text{VII.3})$$

Ist W ein beliebiger Teilraum von V , sodass $\beta|_W$ nicht-degeneriert ist, dann muss $W \cap \ker(\beta) = \{0\}$ gelten, und daher $\dim(W) \leq \dim(V) - \dim \ker(\beta) = \text{rank}(\beta)$, siehe (VII.2). Somit gilt in (VII.3) auch die umgekehrte Ungleichheit. \square

VII.1.18. DEFINITION (Quadratische Formen). Ist $\beta: V \times V \rightarrow \mathbb{K}$ eine symmetrische Bilinearform, dann wird die Abbildung

$$q: V \rightarrow \mathbb{K}, \quad q(v) := \beta(v, v),$$

als die mit β assoziierte *quadratische Form* bezeichnet.

Beachte, dass die quadratische Form einer symmetrischen Bilinearform homogen vom Grad zwei ist, d.h.

$$q(\lambda v) = \lambda^2 q(v), \quad \lambda \in \mathbb{K}, \quad v \in V.$$

Gilt $0 \neq 2 \in \mathbb{K}$, dann kann die Bilinearform β aus der quadratischen Form q mit Hilfe der sogenannten *Polarisierungsidentität* zurückgewonnen werden, siehe Proposition VII.1.22 unten. Wir haben die Bedingung $0 \neq 2 \in \mathbb{K}$ schon öfters angetroffen und wollen diese nun formalisieren.

VII.1.19. DEFINITION (Charakteristik eines Körpers). Unter der *Charakteristik* eines Körpers \mathbb{K} verstehen wir die kleinste Zahl $k \in \mathbb{N}$, sodass

$$k = \underbrace{1 + \cdots + 1}_{k \text{ Summanden}} = 0 \in \mathbb{K}.$$

Existiert keine solche Zahl k , d.h. ist $0 \neq k \in \mathbb{K}$, für jedes $k \in \mathbb{N}$, dann wird die Charakteristik von \mathbb{K} als 0 definiert. Wir werden die Charakteristik von \mathbb{K} mit $\text{char}(\mathbb{K})$ bezeichnen.

VII.1.20. BEISPIEL. Es gilt $\text{char}(\mathbb{Q}) = \text{char}(\mathbb{R}) = \text{char}(\mathbb{C}) = 0$ und $\text{char}(\mathbb{Z}_p) = p$, für jede Primzahl p . In einem Körper \mathbb{K} gilt $0 \neq 2 \in \mathbb{K}$ genau dann, wenn $\text{char}(\mathbb{K}) \neq 2$.

VII.1.21. BEMERKUNG. Ist die Charakteristik eines Körpers nicht 0, so muss sie eine Primzahl sein. Ist nämlich $\text{char}(\mathbb{K}) = nm$ mit $n, m \in \mathbb{N}$, dann gilt $0 = nm \in \mathbb{K}$, also o.B.d.A. $0 = n \in \mathbb{K}$ und daher $n = \text{char}(\mathbb{K})$, $m = 1$. Dies zeigt, dass $\text{char}(\mathbb{K})$ keine echten Teiler besitzt und daher eine Primzahl ist.

VII.1.22. PROPOSITION (Polarisierungsformel). Sei V ein Vektorraum über einem Körper mit $\text{char}(\mathbb{K}) \neq 2$ und $\beta: V \times V \rightarrow \mathbb{K}$ eine symmetrische Bilinearform mit assoziierter quadratischer Form $q(v) = \beta(v, v)$. Dann gilt

$$\beta(v, w) = \frac{1}{2}(q(v+w) - q(v) - q(w)) = \frac{1}{4}(q(v+w) - q(v-w)),$$

d.h. β kann aus q zurückgewonnen werden.

BEWEIS. Aus der Bilinearität und Symmetrie von β folgt

$$\beta(v+w, v+w) = \beta(v, v) + 2\beta(v, w) + \beta(w, w),$$

also

$$2\beta(v, w) = q(v+w) - q(v) - q(w).$$

Nach Voraussetzung ist $0 \neq 2 \in \mathbb{K}$, und wir erhalten die erste Polarisierungsformel. Ersetzen wir in der letzten Gleichung w durch $-w$, erhalten wir

$$-2\beta(v, w) = q(v-w) - q(v) - q(w),$$

denn $\beta(v, -w) = -\beta(v, w)$ und $q(-w) = q(w)$. Subtraktion der letzten beiden Gleichungen liefert

$$4\beta(v, w) = q(v+w) - q(v-w),$$

und somit auch die zweite Polarisierungsformel. Beachte, dass wegen $\text{char}(\mathbb{K}) \neq 2$ auch $4 \neq 0 \in \mathbb{K}$. \square

VII.1.23. BEISPIEL. Betrachte den Körper $\mathbb{K} = \mathbb{Z}_2$ und die mit der Matrix $A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ assoziierte symmetrische Bilinearform, $\beta: \mathbb{K}^2 \times \mathbb{K}^2 \rightarrow \mathbb{K}$, $\beta(x, y) = x^t A y = x_1 y_2 + x_2 y_1$. Für die assoziierte quadratische Form gilt $q(x) = \beta(x, x) = x_1 x_2 + x_2 x_1 = 2x_1 x_2 = 0$. In diesem Fall kann die Bilinearform β also *nicht* aus der zugehörigen quadratischen Form rekonstruiert werden. Die Voraussetzung $\text{char}(\mathbb{K}) \neq 2$ in Proposition VII.1.22 kann daher nicht ersatzlos gestrichen werden.

VII.1.24. BEMERKUNG. Sei \mathbb{K} ein Körper mit $\text{char}(\mathbb{K}) \neq 2$ und $\beta: V \times V \rightarrow \mathbb{K}$ eine symmetrische Bilinearform mit assoziierter quadratischer Form $q(v) = \beta(v, v)$. Dann gilt

$$O(V, \beta) = \{\varphi \in \text{GL}(V) \mid \forall v \in V : q(\varphi(v)) = q(v)\}.$$

Dies folgt sofort aus der Polarisierungsidentität in Proposition VII.1.22 und der Linearität von φ , siehe Bemerkung VII.1.11 für die Definition von $O(V, \beta)$.

VII.1.25. SATZ. Sei V ein endlich dimensionaler Vektorraum über einem Körper \mathbb{K} mit $\text{char}(\mathbb{K}) \neq 2$. Weiters sei $\beta: V \times V \rightarrow \mathbb{K}$ eine symmetrische Bilinearform. Dann existieren eine geordnete Basis B von V , sodass $[\beta]_B$ Diagonalgestalt hat. Ist \mathbb{K} ein Körper indem jedes Element mindestens eine Quadratwurzel besitzt (etwa jeder algebraisch abgeschlossene Körper), dann kann B so gewählt werden, dass $[\beta]_B = \begin{pmatrix} I_k & 0 \\ 0 & 0 \end{pmatrix}$, wobei $k = \text{rank}(\beta)$.

BEWEIS. Wir führen den Beweis mittels Induktion nach der Dimension von V . Der Induktionsanfang, $\dim(V) = 1$, ist trivial. Nun zum Induktionsschritt. Ist $\beta = 0$, so hat jede Basis von V die gewünschte Eigenschaft. O.B.d.A. sei also $\beta \neq 0$. Nach Proposition VII.1.22 existiert $b_1 \in V$, sodass $a_1 := \beta(b_1, b_1) \neq 0$. Beachte, dass

$$W := \{v \in V \mid \beta(b_1, v) = 0\} = \ker(\beta(b_1, -): V \rightarrow \mathbb{K})$$

eine Hyperebene in V ist, die b_1 nicht enthält. Somit $V = \langle b_1 \rangle \oplus W$. Wenden wir die Induktionsvoraussetzung auf die symmetrische Bilinearform $\beta|_W$ an, erhalten wir eine Basis b_2, \dots, b_n von W , sodass $\beta(b_i, b_j) = a_i \delta_{ij}$, $2 \leq i, j \leq n$. Beachte, dass dies auch für $i = 1$ oder $j = 1$ richtig bleibt, $\beta(b_i, b_j) = a_i \delta_{ij}$, $1 \leq i, j \leq n$. Somit ist $B = (b_1, \dots, b_n)$ eine Basis von V und $[\beta]_B$ hat Diagonalgestalt. Dies zeigt den ersten Teil des Satzes.

Sei nun \mathbb{K} ein Körper in dem jedes Element eine Quadratwurzel besitzt und $k := \text{rank}(\beta)$. Durch Umnummerieren der Basisvektoren können wir $a_i \neq 0$ für $i = 1, \dots, k$ und $a_{k+1} = \dots = a_n = 0$ erreichen. Setzen wir nun

$$\tilde{b}_i := \begin{cases} \frac{1}{\sqrt{a_i}} b_i & \text{für } i = 1, \dots, k \\ b_i & \text{für } i = k+1, \dots, n \end{cases}$$

dann ist $\tilde{B} := (\tilde{b}_1, \dots, \tilde{b}_n)$ eine Basis von V und $[\beta]_{\tilde{B}} = \begin{pmatrix} I_k & 0 \\ 0 & 0 \end{pmatrix}$. \square

VII.1.26. KOROLLAR. Sei V ein n -dimensionaler komplexer Vektorraum und $\beta: V \times V \rightarrow \mathbb{C}$ eine nicht-degenerierte symmetrische Bilinearform. Dann existiert eine Basis B von V , sodass $[\beta]_B = I_n$, d.h.

$$\beta(v, w) = [v]_B^t [w]_B, \quad v, w \in V.$$

Jede solche Basis liefert einen Gruppenisomorphismus,

$$O(V, \beta) \cong O_n(\mathbb{C}) := \{A \in M_{n \times n}(\mathbb{C}) \mid A^t A = I_n\}, \quad \varphi \leftrightarrow [\varphi]_{BB}.$$

BEWEIS. Die erste Behauptung folgt sofort aus Satz VII.1.25. Die zweite Aussage erhalten wir aus den Bemerkungen VII.1.11 und VII.1.16. \square

Zwei symmetrische Matrizen $A, A' \in M_{n \times n}^{\text{sym}}(\mathbb{K})$ werden *kongruent* genannt, falls $S \in \text{GL}_n(\mathbb{K})$ existiert, sodass $S^t A S = A'$. Dies definiert eine Äquivalenzrelation auf der Menge der symmetrischen Matrizen, $M_{n \times n}^{\text{sym}}(\mathbb{K}) := \{A \in M_{n \times n}(\mathbb{K}) : A^t = A\}$, siehe Aufgabe 87. Ist \mathbb{K} ein algebraisch abgeschlossener Körper (oder, allgemeiner, ein Körper in dem jedes Element eine Quadratwurzel besitzt) mit $\text{char}(\mathbb{K}) \neq 2$ und ist $A \in M_{n \times n}^{\text{sym}}(\mathbb{K})$ eine symmetrische Matrix, dann existiert eine invertierbare Matrix $S \in \text{GL}_n(\mathbb{K})$, sodass $S^t A S = \begin{pmatrix} I_k & 0 \\ 0 & 0 \end{pmatrix}$, wobei $k = \text{rank}(A)$. Dies folgt sofort aus Satz VII.1.25 durch Betrachten der Bilinearform $\beta_A: \mathbb{K}^n \times \mathbb{K}^n \rightarrow \mathbb{K}$, $\beta_A(x, y) = x^t A y$. Aus diesen Erläuterungen folgt, dass zwei symmetrische Matrizen über einem Körper \mathbb{K} wie oben, genau dann kongruent sind, wenn sie gleichen Rang haben. Insbesondere erhalten wir:

VII.1.27. KOROLLAR. *Zwei symmetrische komplexe $(n \times n)$ -Matrizen sind genau dann kongruent wenn sie gleichen Rang haben.*

VII.1.28. BEISPIEL. Betrachte die symmetrische Matrix

$$A = \begin{pmatrix} 1 & -2 & -3 \\ -2 & 5 & 7 \\ -3 & 7 & 11 \end{pmatrix} \in M_{3 \times 3}^{\text{sym}}(\mathbb{C})$$

Nach Korollar VII.1.26 existiert eine Basis B von \mathbb{C}^3 , sodass $[\beta_A]_B = I_3$, denn $\text{rank}(A) = 3$. Um eine solche Basis zu bestimmen gehen wir wie im Beweis von Satz VII.1.25 vor. Da $\beta_A(e_1, e_1) = e_1^t A e_1 = 1 \neq 0$ können wir $b_1 = e_1$ als ersten Basisvektor verwenden,

$$b_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad \beta_A(b_1, b_1) = b_1^t A b_1 = 1.$$

Es ist nun b_2 so zu bestimmen, dass $\beta_A(b_1, b_2) = 0$ und $\beta_A(b_2, b_2) = 1$. Wir entscheiden uns für

$$b_2 = \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix}, \quad \beta_A(b_1, b_2) = b_1^t A b_2 = 0, \quad \beta_A(b_2, b_2) = b_2^t A b_2 = 1.$$

Für den letzten Basisvektor, b_3 , muss $\beta_A(b_1, b_3) = \beta_A(b_2, b_3) = 0$ und $\beta_A(b_3, b_3) = 1$ gelten. Wir verwenden

$$b_3 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \quad \beta_A(b_1, b_3) = 0, \quad \beta_A(b_2, b_3) = 0, \quad \beta_A(b_3, b_3) = 1.$$

Nach Konstruktion ist $B = (b_1, b_2, b_3)$ eine Basis von \mathbb{C}^3 mit $[\beta_A]_B = I_3$. Für die Basiswechsellmatrix,

$$S := T_{EB} = \begin{pmatrix} 1 & 2 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad \text{gilt daher} \quad S^t A S = T_{EB}^t [\beta_A]_E T_{EB} = [\beta_A]_B = I_3,$$

wobei $E = (e_1, e_2, e_3)$ die Standardbasis von \mathbb{C}^3 bezeichnet. Alternativ, können wir die assoziierte quadratische Form,

$$q_A \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x \\ y \\ z \end{pmatrix}^t \begin{pmatrix} 1 & -2 & -3 \\ -2 & 5 & 7 \\ -3 & 7 & 11 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = x^2 + 5y^2 + 11z^2 - 4xy - 6xz + 14yz,$$

durch Ergänzen auf vollständige Quadrate auf die Form

$$q_A \begin{pmatrix} x \\ y \\ z \end{pmatrix} = (x - 2y - 3z)^2 + y^2 + 2z^2 + 2yz = (x - 2y - 3z)^2 + (y + z)^2 + z^2$$

bringen. Dies zeigt $q_A(v) = (Tv)^tTv = v^tT^tTv$ für alle $v \in \mathbb{C}^3$, wobei

$$T := \begin{pmatrix} 1 & -2 & -3 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}.$$

Durch Polarisieren erhalten wir $v^tAw = \beta_A(v, w) = v^tT^tTw$, also $A = T^tT$ oder $(T^{-1})^tAT^{-1} = I_3$. Tatsächlich gilt, aufgrund unserer Wahlen, $T^{-1} = S$.

Über dem Körper $\mathbb{K} = \mathbb{R}$ ist die Situation ein wenig komplizierter.

VII.1.29. DEFINITION (Definitheit). Eine symmetrische Bilinearform auf einem reellen Vektorraum, $\beta: V \times V \rightarrow \mathbb{R}$, und die damit assoziierte quadratische Form, $q: V \rightarrow \mathbb{R}$, $q(v) = \beta(v, v)$, heißen:

- (a) *positiv semidefinit*, in Zeichen $\beta \geq 0$, falls $\beta(v, v) \geq 0$, für alle $v \in V$.
- (b) *positiv definit*, in Zeichen $\beta > 0$, falls $\beta(v, v) > 0$, für alle $0 \neq v \in V$.
- (c) *negativ semidefinit*, in Zeichen $\beta \leq 0$, falls $\beta(v, v) \leq 0$, für alle $v \in V$.
- (d) *negativ definit*, in Zeichen $\beta < 0$, falls $\beta(v, v) < 0$, für alle $0 \neq v \in V$.
- (e) *indefinit*, falls es $v, w \in V$ gibt, sodass $\beta(v, v) > 0$ und $\beta(w, w) < 0$.

Eine symmetrische Matrix $A \in M_{n \times n}(\mathbb{R})$ wird positiv bzw. negativ (semi) definit genannt, falls die damit assoziierte symmetrische Bilinearform $\beta_A: \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$, $\beta_A(x, y) = x^tAy$, positiv bzw. negativ (semi)definit ist, d.h.: $A \geq 0$ falls $x^tAx \geq 0$ für alle $x \in \mathbb{R}^n$; $A > 0$ falls $x^tAx > 0$ für alle $0 \neq x \in \mathbb{R}^n$; $A \leq 0$ falls $x^tAx \leq 0$ für alle $x \in \mathbb{R}^n$; und $A < 0$ falls $x^tAx < 0$ für alle $0 \neq x \in \mathbb{R}^n$.

VII.1.30. BEMERKUNG. Beachte, dass eine symmetrische Bilinearform β genau dann *negativ (semi)definit* ist, wenn die symmetrische Bilinearform $-\beta$ positiv (semi)definit ist. Jede positiv definite symmetrische Bilinearform hat trivialen Kern und ist daher nicht-degeneriert. Dasselbe gilt für negativ definite symmetrische Bilinearformen. Ist β eine positiv definite symmetrische Bilinearform auf V und $W \subseteq V$ ein Teilraum, dann ist auch die Einschränkung $\beta|_W$ positiv definit und daher nicht-degeneriert. Beachte, dass die Einschränkung einer nicht-degenerierten symmetrischen Bilinearform sehr wohl degeneriert sein kann. Schränken wir etwa die Lorentz Metrik g aus Beispiel VII.1.5 auf den 1-dimensionalen Teilraum $W := \langle e_0 + e_1 \rangle \subseteq \mathbb{R}^4$ ein, so erhalten wir $g|_W = 0$. Nicht-triviale Teilräume $W \neq \{0\}$ mit $\beta|_W = 0$ werden *isotrop* genannt.

VII.1.31. BEMERKUNG (Konvexität). Seien β und β' zwei positiv definite symmetrische Bilinearformen auf einem reellen Vektorraum V . Für $0 \leq \lambda \leq 1$ ist

dann auch $\lambda\beta + (1 - \lambda)\beta'$ eine positiv definite symmetrische Bilinearform auf V , denn für jedes $0 \neq v \in V$ gilt:

$$(\lambda\beta + (1 - \lambda)\beta')(v, v) = \lambda \underbrace{\beta(v, v)}_{>0} + (1 - \lambda) \underbrace{\beta'(v, v)}_{>0} > 0.$$

Die Menge der positiv definiten symmetrischen Bilinearformen bildet daher eine konvexe⁶ Teilmenge im Vektorraum aller symmetrischen Bilinearformen auf V . Auch die Menge der positiv semidefiniten symmetrischen Bilinearformen ist eine konvexe Teilmenge. Analoge Aussagen gelten für negativ (semi)definite symmetrische Bilinearformen. Beachte, dass die Menge der nicht-degenerierten symmetrischen Bilinearformen nicht konvex ist.

VII.1.32. BEISPIEL. Das Euklidische innere Produkt aus Beispiel VII.1.4 ist positiv definit, denn für jedes $0 \neq x \in \mathbb{R}^n$ gilt

$$\langle x, x \rangle = x_1^2 + \cdots + x_n^2 > 0.$$

Auch die symmetrische Bilinearform aus Beispiel VII.1.6 ist positiv definit, denn für jedes $0 \neq A \in M_{m \times n}(\mathbb{R})$ gilt

$$\text{tr}(A^t A) = \sum_{j=1}^n \sum_{i=1}^m A_{ij}^2 > 0.$$

Auch die symmetrische Bilinearform aus Beispiel VII.1.7 ist positiv definit, denn für jedes $0 \neq f \in C^0(I, \mathbb{R})$ gilt

$$\int_I f^2(t) dt > 0.$$

Die symmetrische Bilinearform $\beta(x, y) = x^t \begin{pmatrix} I_p & 0 \\ 0 & 0 \end{pmatrix} y$, ist positiv semidefinit, denn für jedes $x \in \mathbb{R}^n$ gilt

$$\beta(x, x) = x^t \begin{pmatrix} I_p & 0 \\ 0 & 0 \end{pmatrix} x = x_1^2 + \cdots + x_p^2 \geq 0.$$

Die Lorentz Metrik aus Beispiel VII.1.5 ist indefinit, denn $g(e_0, e_0) < 0$ und $g(e_1, e_1) > 0$.

VII.1.33. BEISPIEL. Die symmetrische Bilinearform $\beta(x, y) = x^t \begin{pmatrix} 1 & -1 \\ -1 & 2 \end{pmatrix} y$ ist positiv definit, denn

$$q \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \beta \left(\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \right) = x_1^2 - 2x_1x_2 + 2x_2^2 = (x_1 - x_2)^2 + x_2^2 > 0,$$

für alle $0 \neq \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \in \mathbb{R}^2$.

⁶Eine Teilmenge A eines Vektorraums wird konvex genannt, wenn sie folgende Eigenschaft besitzt: $\forall x, y \in A \forall \lambda \in [0, 1] : \lambda x + (1 - \lambda)y \in A$. Dies bedeutet, dass für je zwei Punkte x und y in A auch die Strecke von x nach y zur Gänze in A liegt.

VII.1.34. BEISPIEL. Die symmetrische Bilinearform $\beta(x, y) = x^t \begin{pmatrix} 1 & -2 \\ -2 & 2 \end{pmatrix} y$ ist indefinit, denn

$$q \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \beta \left(\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \right) = x_1^2 - 4x_1x_2 + 2x_2^2 = (x_1 - 2x_2)^2 - 2x_2^2,$$

also $\beta \left(\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \end{pmatrix} \right) = 1 > 0$ und $\beta \left(\begin{pmatrix} 2 \\ 1 \end{pmatrix}, \begin{pmatrix} 2 \\ 1 \end{pmatrix} \right) = -2 < 0$.

VII.1.35. SATZ (Trägheitssatz von Sylvester). *Sei V ein n -dimensionaler reeller Vektorraum und $\beta: V \times V \rightarrow \mathbb{R}$ eine symmetrische Bilinearform. Dann existiert eine geordnete Basis B von V , sodass*

$$[\beta]_B = \begin{pmatrix} I_p & & \\ & -I_q & \\ & & 0 \end{pmatrix}.$$

Dabei ist $p + q = \text{rank}(\beta)$ und

$$p = \max \{ \dim(W) \mid W \text{ Teilraum von } V \text{ mit } \beta|_W > 0 \}, \quad (\text{VII.4})$$

$$q = \max \{ \dim(W) \mid W \text{ Teilraum von } V \text{ mit } \beta|_W < 0 \}. \quad (\text{VII.5})$$

Insbesondere sind die Zahlen p und q unabhängig von der Basis B .

BEWEIS. Nach Satz VII.1.25 existiert eine Basis $\tilde{B} = (\tilde{b}_1, \dots, \tilde{b}_n)$ von V , sodass $\beta(\tilde{b}_i, \tilde{b}_j) = a_i \delta_{ij}$, wobei $a_1, \dots, a_p > 0$, $a_{p+1}, \dots, a_{p+q} < 0$ und $a_{p+q+1} = \dots = a_n = 0$. Setzen wir

$$b_i := \begin{cases} \frac{1}{\sqrt{a_i}} \tilde{b}_i & \text{für } i = 1, \dots, p, \\ \frac{1}{\sqrt{-a_i}} \tilde{b}_i & \text{für } i = p+1, \dots, p+q, \\ \tilde{b}_i & \text{für } i = p+q+1, \dots, n, \end{cases}$$

so ist $B = (b_1, \dots, b_n)$ eine Basis von V mit der gewünschten Eigenschaft,

$$[\beta]_B = \begin{pmatrix} I_p & & \\ & -I_q & \\ & & 0 \end{pmatrix}.$$

Betrachten wir die Teilräume $V' := \langle b_1, \dots, b_p \rangle$ und $V'' := \langle b_{p+1}, \dots, b_n \rangle$, so gilt

$$V = V' \oplus V'', \quad \beta|_{V'} > 0 \quad \text{und} \quad \beta|_{V''} \leq 0.$$

Da $\dim(V') = p$ erhalten wir insbesondere

$$p \leq \max \{ \dim(W) \mid W \text{ Teilraum von } V \text{ mit } \beta|_W > 0 \}. \quad (\text{VII.6})$$

Sei nun W ein beliebiger Teilraum von V mit $\beta|_W > 0$. Dann gilt

$$W \cap V'' = \{0\},$$

denn für $v \in W \cap V''$ ist $\beta(v, v) = 0$ und daher $v = 0$. Daraus erhalten wir

$$\dim(W) + \dim(V'') = \dim(W + V'') \leq \dim(V) = n,$$

also $\dim(W) \leq n - \dim(V'') = p$. Zusammen mit (VII.6) erhalten wir (VII.4). Die Formel (VII.5) lässt sich analog zeigen, oder aus (VII.4) für $-\beta$ ablesen. \square

VII.1.36. DEFINITION (Signatur). Das Paar (p, q) in Satz VII.1.35 wird als *Signatur* der symmetrischen Bilinearform β bezeichnet.⁷ Unter der Signatur einer symmetrischen Matrix A bzw. der Signatur einer quadratischen Form verstehen wir die Signatur der damit assoziierten symmetrischen Bilinearform.

VII.1.37. BEISPIEL. Das Euklidische innere Produkt aus Beispiel VII.1.4 hat Signatur $(n, 0)$. Die Lorentz Metrik aus Beispiel VII.1.5 hat Signatur $(3, -1)$. Die symmetrische Matrix

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

hat Signatur (p, p) , denn

$$S^t A S = \begin{pmatrix} I_p & 0 \\ 0 & -I_p \end{pmatrix} \quad \text{mit} \quad S = \frac{1}{\sqrt{2}} \begin{pmatrix} I_p & -I_p \\ I_p & I_p \end{pmatrix}.$$

VII.1.38. KOROLLAR. Zwei reelle symmetrische $(n \times n)$ -Matrizen sind genau dann kongruent, wenn sie gleiche Signatur haben.

VII.1.39. KOROLLAR. Sei $\beta: V \times V \rightarrow \mathbb{R}$ eine positiv definite symmetrische Bilinearform auf einem n -dimensionalen reellen Vektorraum V . Dann existiert eine Basis B von V , sodass $[\beta]_B = I_n$, d.h.

$$\beta(v, w) = [v]_B^t [w]_B, \quad v, w \in V.$$

Jede solche Basis liefert einen Gruppenisomorphismus,

$$O(V, \beta) \cong O_n := \{A \in M_{n \times n}(\mathbb{R}) \mid A^t A = I_n\}, \quad \varphi \leftrightarrow [\varphi]_{BB}.$$

BEWEIS. Die erste Behauptung folgt sofort aus Satz VII.1.35. Die zweite Aussage erhalten wir aus den Bemerkungen VII.1.11 und VII.1.16. \square

VII.1.40. BEISPIEL. Wir wollen die Signatur der mit symmetrischen Matrix

$$A = \begin{pmatrix} 4 & 2 & -2 \\ 2 & -3 & -7 \\ -2 & -7 & -8 \end{pmatrix}$$

assoziierten symmetrischen Bilinearform $\beta_A(v, w) = v^t A w$ auf \mathbb{R}^3 bestimmen. Durch Ergänzen auf vollständige Quadrate erhalten wir

$$\begin{aligned} q \begin{pmatrix} x \\ y \\ z \end{pmatrix} &= 4x^2 - 3y^2 - 8z^2 + 4xy - 4xz - 14yz \\ &= (2x + y - z)^2 - 4y^2 - 9z^2 - 12yz = (2x + y - z)^2 - (2y + 3z)^2. \end{aligned}$$

Die Signatur von β_A ist daher $(1, 1)$. Obige Rechnung zeigt,

$$v^t A v = q(v) = (Tv)^t \begin{pmatrix} 1 & & \\ & -1 & \\ & & 0 \end{pmatrix} (Tv), \quad \text{mit} \quad T = \begin{pmatrix} 2 & 1 & -1 \\ 0 & 2 & 3 \\ 0 & 0 & 1 \end{pmatrix},$$

also

$$A = T^t \begin{pmatrix} 1 & & \\ & -1 & \\ & & 0 \end{pmatrix} T,$$

⁷Manchmal wird auch die Differenz, $p - q$, als *Signatur* von β bezeichnet.

und daher

$$S^t A S = \begin{pmatrix} 1 & & \\ & -1 & \\ & & 0 \end{pmatrix}, \quad \text{wobei} \quad S = T^{-1} = \begin{pmatrix} \frac{1}{2} & -\frac{1}{4} & \frac{5}{4} \\ 0 & \frac{1}{2} & -\frac{3}{2} \\ 0 & 0 & 1 \end{pmatrix}.$$

VII.1.41. DEFINITION (Hauptminoren). Unter dem k -ten *Hauptminor* einer symmetrischen Matrix $A \in M_{n \times n}^{\text{sym}}(\mathbb{R})$, verstehen wir den Skalar

$$\det \begin{pmatrix} a_{11} & \cdots & a_{1k} \\ \vdots & & \vdots \\ a_{k1} & \cdots & a_{kk} \end{pmatrix},$$

wobei a_{ij} die Eintragung der Matrix A bezeichnen, $1 \leq k \leq n$.

VII.1.42. SATZ (Sylvester Kriterium). *Sei V ein endlich dimensionaler reeller Vektorraum. Eine symmetrische Bilinearform $\beta: V \times V \rightarrow \mathbb{R}$ ist genau dann positiv definit, wenn alle Hauptminoren der Matrix $[\beta]_B$ positiv sind, bezüglich einer (und dann jeder) Basis B von V .*

BEWEIS. Wir zeigen zunächst, dass eine positiv definite symmetrische Bilinearform positive Hauptminoren hat. Sei dazu β eine positiv definite symmetrische Bilinearform auf V . Nach Satz VII.1.35 existiert eine Basis $B = (b_1, \dots, b_n)$ von V , sodass $[\beta]_B = I_n$. Es gilt daher $\det([\beta]_B) > 0$. Ist C eine weitere Basis von V , und bezeichnet $S = T_{CB} \in \text{GL}_n(\mathbb{R})$ die Basiswechselmatrix, dann haben wir $[\beta]_C = S^t [\beta]_B S$, siehe Proposition VII.1.9, also auch

$$\det([\beta]_C) = \det(S^t [\beta]_B S) = \det(S^t) \det([\beta]_B) \det(S) = \det(S)^2 \det([\beta]_B) > 0,$$

denn $\det(S) \neq 0$. Dies zeigt $\det([\beta]_B) > 0$, für jede Basis B von V . Sei nun $k \in \{1, \dots, n\}$ und bezeichne $W := \langle b_1, \dots, b_k \rangle$ den von den ersten k Basisvektoren aufgespannten Teilraum. Dann ist $\beta|_W$ positiv definit, $\tilde{B} = (b_1, \dots, b_k)$ ist eine Basis von W und aus dem eben Gezeigten folgt $\det([\beta|_W]_{\tilde{B}}) > 0$. Beachte, dass $\det([\beta|_W]_{\tilde{B}})$ gerade der k -te Hauptminor von $[\beta]_B$ ist, denn

$$[\beta]_B = \begin{pmatrix} [\beta|_W]_{\tilde{B}} & * \\ * & * \end{pmatrix}.$$

Somit sind also alle Hauptminoren von $[\beta]_B$ positiv.

Für die umgekehrte Implikation sei nun B eine Basis von V , sodass alle Hauptminoren der Matrix $[\beta]_B$ positiv sind. Wir werden nun mittels Induktion nach $\dim(V)$ zeigen, dass β positiv definit ist. Der Induktionsanfang, $\dim(V) = 1$, ist trivial. Für den Induktionsschritt betrachten wir die Basis $C = (c_1, \dots, c_n)$ von V , wobei

$$c_1 := b_1, \quad c_i := b_i - \frac{\beta(b_i, b_1)}{\beta(b_1, b_1)} b_1, \quad i = 2, \dots, n. \quad (\text{VII.7})$$

Beachte, dass $\beta(b_1, b_1) > 0$, da der erste Hauptminor von $[\beta]_B$ positiv ist. Die Transformationsmatrix zum Basiswechsel hat die Gestalt

$$S = T_{BC} = \begin{pmatrix} 1 & * & \cdots & * \\ & 1 & \ddots & \vdots \\ & & \ddots & * \\ & & & 1 \end{pmatrix}.$$

Ist X eine Matrix so schreiben wir $X_{(k)}$ für die linke obere $(k \times k)$ -Untermatrix von X , d.h. der k -te Hauptminor von X ist $\det(X_{(k)})$. Aus $[\beta]_C = S^t[\beta]_B S$ und der speziellen Gestalt von S folgt

$$([\beta]_C)_{(k)} = S_{(k)}^t([\beta]_B)_{(k)} S_{(k)},$$

und daher

$$\det\left([\beta]_C\right)_{(k)} = \underbrace{\det(S_{(k)}^t)}_{=1} \det\left([\beta]_B\right)_{(k)} \underbrace{\det(S_{(k)})}_{=1} = \det\left([\beta]_B\right)_{(k)}.$$

Dies zeigt, dass $[\beta]_C$ die selben Hauptminoren wie $[\beta]_B$ hat. Insbesondere sind auch alle Hauptminoren von $[\beta]_C$ positiv. Wir betrachten nun den Teilraum $W := \langle c_2, \dots, c_n \rangle$ mit Basis $\tilde{C} := (c_2, \dots, c_n)$. Nach Konstruktion der Basis C , siehe (VII.7), gilt

$$[\beta]_C = \begin{pmatrix} \beta(b_1, b_1) & 0 \\ 0 & [\beta|_W]_{\tilde{C}} \end{pmatrix}.$$

Da $\beta(b_1, b_1) > 0$ sind auch alle Hauptminoren von $[\beta|_W]_{\tilde{C}}$ positiv. Nach Induktionsvoraussetzung ist daher $\beta|_W > 0$. Ist nun $v \in V$ beliebig, dann existieren $\lambda \in \mathbb{K}$ und $w \in W$, sodass $v = \lambda b_1 + w$ und wir erhalten

$$\beta(v, v) = \beta(\lambda b_1 + w, \lambda b_1 + w) = \lambda^2 \underbrace{\beta(b_1, b_1)}_{>0} + 2\lambda \underbrace{\beta(b_1, w)}_{=0} + \underbrace{\beta(w, w)}_{\geq 0} \geq 0$$

wobei Gleichheit nur eintreten kann, wenn $\lambda = 0$ und $\beta(w, w) = 0$, d.h. nur wenn $v = 0$. Dies zeigt, dass β positiv definit ist. \square

VII.1.43. KOROLLAR. Sei V ein n -dimensionaler reeller Vektorraum. Eine symmetrische Bilinearform $\beta: V \times V \rightarrow \mathbb{R}$ ist genau dann negativ definit, wenn alle Hauptminoren der Matrix $-[\beta]_B$ positiv sind, bezüglich einer (und dann jeder) Basis B von V . Dies ist genau dann der Fall, wenn

$$(-1)^k \det \begin{pmatrix} a_{11} & \cdots & a_{1k} \\ \vdots & & \vdots \\ a_{k1} & \cdots & a_{kk} \end{pmatrix} > 0,$$

für jedes $k = 1, \dots, n$, wobei a_{ij} die Eintragungen der Matrix $[\beta]_B$ bezeichnen.

BEWEIS. Dies folgt aus Satz VII.1.35, da $\beta < 0 \Leftrightarrow -\beta > 0$ und $[-\beta]_B = -[\beta]_B$. Beachte auch $\det(-C) = (-1)^k \det(C)$, für jede $(k \times k)$ -Matrix C . \square

VII.1.44. BEISPIEL. Betrachte die symmetrische Matrix

$$A = \begin{pmatrix} -1 & 2 & 3 \\ 2 & -5 & -7 \\ 3 & -7 & -11 \end{pmatrix} \in M_{3 \times 3}^{\text{sym}}(\mathbb{R}).$$

Für ihre Hauptminoren erhalten wir

$$\det(-1) = -1 < 0, \quad \det \begin{pmatrix} -1 & 2 \\ 2 & -5 \end{pmatrix} = 3 > 0, \quad \det \begin{pmatrix} -1 & 2 & 3 \\ 2 & -5 & -7 \\ 3 & -7 & -11 \end{pmatrix} = -1 < 0,$$

also ist die Matrix A nach Korollar VII.1.43 negativ definit. Alternativ können wir die assoziierte quadratische Form,

$$q \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x \\ y \\ z \end{pmatrix}^t \begin{pmatrix} -1 & 2 & 3 \\ 2 & -5 & -7 \\ 3 & -7 & -11 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = -x^2 - 5y^2 - 11z^2 + 4xy + 6xz - 14yz,$$

auf vollständige Quadrate ergänzen,

$$q \begin{pmatrix} x \\ y \\ z \end{pmatrix} = -(x - 2y - 3z)^2 - y^2 - 2z^2 - 2yz = -(x - 2y - 3z)^2 - (y + z)^2 - z^2,$$

und daraus schließen, dass sie negativ definit ist.

VII.1.45. BEMERKUNG (Lokale Extrema). Sei $U \subseteq \mathbb{R}^n$ eine offene Teilmenge, $f: U \rightarrow \mathbb{R}$ eine C^2 -Funktion und $y \in U$ ein kritischer Punkt von f , d.h.

$$D_y f = \left(\frac{\partial f}{\partial x_1}(y), \dots, \frac{\partial f}{\partial x_n}(y) \right) = 0.$$

Da $\frac{\partial^2 f}{\partial x_i \partial x_j} = \frac{\partial^2 f}{\partial x_j \partial x_i}$ ist die Hessesche Matrix,

$$H_y f = \begin{pmatrix} \frac{\partial^2 f}{\partial x_1 \partial x_1}(y) & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_n}(y) \\ \vdots & & \vdots \\ \frac{\partial^2 f}{\partial x_n \partial x_1}(y) & \cdots & \frac{\partial^2 f}{\partial x_n \partial x_n}(y) \end{pmatrix},$$

symmetrisch. Aus der Analysisvorlesung wissen wir: Ist die Hessesche positiv definit, $H_y f > 0$, dann ist y ein lokales Minimum von f . Analog folgt aus $H_y f < 0$, dass y ein lokales Maximum von f ist.

VII.1.46. DEFINITION (Sesquilinearformen). Sei V ein komplexer Vektorraum. Unter einer *Sesquilinearform* auf V verstehen wir eine Abbildung

$$h: V \times V \rightarrow \mathbb{C},$$

die komplex linear in der zweiten Eintragung und komplex anti-linear in der ersten Eintragung ist, d.h. es gilt

$$h(v, w + w') = h(v, w) + h(v, w'), \quad h(v, \lambda w) = \lambda h(v, w),$$

und

$$h(v + v', w) = h(v, w) + h(v', w), \quad h(\lambda v, w) = \bar{\lambda} h(v, w),$$

für alle $v, v', w, w' \in V$ und $\lambda \in \mathbb{C}$. Unter einer *Hermiteischen Form* auf V verstehen wir eine Sesquilinearform h auf V , die *symmetrisch* in folgendem Sinn ist: für alle $v, w \in V$ gilt:

$$h(w, v) = \overline{h(v, w)}.$$

VII.1.47. BEMERKUNG. Die Menge aller Sesquilinearformen auf V bildet einen komplexen Vektorraum bezüglich der Operationen

$$(h + h')(v, w) := h(v, w) + h'(v, w) \quad \text{und} \quad (\lambda h)(v, w) := \lambda h(v, w).$$

Dabei bezeichnen h und h' zwei Sesquilinearformen auf V , $v, w \in V$ und $\lambda \in \mathbb{C}$. Die Menge der Hermiteschen Formen ist *kein* komplexer Teilraum, denn für eine Hermitesche Form, h , ist $\tilde{h} := \mathbf{i}h$ *keine* Hermitesche Form, da ja $\tilde{h}(w, v) = \mathbf{i}h(w, v) = \overline{\mathbf{i}h(v, w)} = \overline{-\tilde{h}(v, w)}$. Die Hermiteschen Formen bilden jedoch einen *reellen* Teilraum, d.h. für Hermitesche Formen h, h' auf V und $\lambda \in \mathbb{R}$ sind auch $h + h'$ sowie λh Hermitesche Formen auf V .

Ist $A \in M_{m \times n}(\mathbb{C})$ eine komplexe Matrix, dann wird $A^* := \bar{A}^t \in M_{n \times m}(\mathbb{C})$ als die zu A *adjungierte Matrix* bezeichnet, d.h. $(A^*)_{ij} = \overline{A_{ji}}$. Etwa gilt

$$\begin{pmatrix} 1 & 2\mathbf{i} & -7\mathbf{i} \\ 2 & 1-\mathbf{i} & -1-3\mathbf{i} \end{pmatrix}^* = \begin{pmatrix} 1 & 2 \\ -2\mathbf{i} & 1+\mathbf{i} \\ 7\mathbf{i} & -1+3\mathbf{i} \end{pmatrix}.$$

Für die Adjungierte gelten folgende Rechenregeln, siehe Aufgabe 94:

- (a) $(A + B)^* = A^* + B^*$, für alle $A, B \in M_{m \times n}(\mathbb{C})$.
- (b) $(\lambda A)^* = \bar{\lambda}A^*$, für alle $A \in M_{m \times n}(\mathbb{C})$ und $\lambda \in \mathbb{C}$.
- (c) $(AB)^* = B^*A^*$, für alle $A \in M_{m \times n}(\mathbb{C})$ und $B \in M_{n \times l}(\mathbb{C})$.

Eine quadratische Matrix $A \in M_{n \times n}(\mathbb{C})$ wird *Hermitesch* oder *selbstadjungiert* genannt, wenn $A^* = A$. Die Menge der selbstadjungierten Matrizen bildet einen reellen, aber keinen komplexen, Teilraum von $M_{n \times n}(\mathbb{C})$.

VII.1.48. BEISPIEL (Sesquilinearform einer Matrix). Jedes $A \in M_{n \times n}(\mathbb{C})$ definiert eine Sesquilinearform h_A auf \mathbb{C}^n ,

$$h_A: \mathbb{C}^n \times \mathbb{C}^n \rightarrow \mathbb{C}, \quad h_A(x, y) = x^*Ay, \quad x, y \in \mathbb{C}^n,$$

denn $(x_1 + x_2)^*Ay = x_1^*Ay + x_2^*Ay$, $(\lambda x)^*Ay = \bar{\lambda}x^*Ay$, $x^*A(y_1 + y_2) = x^*Ay_1 + x^*Ay_2$ und $x^*A(\lambda y) = \lambda x^*Ay$, für $x, x_1, x_2, y, y_1, y_2 \in \mathbb{C}^n$ und $\lambda \in \mathbb{C}$. Die Zuordnung $A \leftrightarrow h_A$ liefert einen linearen Isomorphismus zwischen $M_{n \times n}(\mathbb{C})$ und dem komplexen Vektorraum aller Sesquilinearformen auf \mathbb{C}^n . Die Sesquilinearform h_A ist genau dann eine Hermitesche Form, wenn $A^* = A$ gilt.

VII.1.49. DEFINITION (Matrixdarstellung). Sei $h: V \times V \rightarrow \mathbb{K}$ eine Sesquilinearform auf einem n -dimensionalen komplexen Vektorraum V . Ist $B = (b_1, \dots, b_n)$ eine geordnete Basis von V , dann wird die Matrix $[h]_B \in M_{n \times n}(\mathbb{C})$,

$$([h]_B)_{ij} := h(b_i, b_j), \quad 1 \leq i, j \leq n,$$

als *Matrixdarstellung von h bezüglich der Basis B* bezeichnet.

VII.1.50. PROPOSITION. Sei V ein n -dimensionaler komplexer Vektorraum. Jede geordnete Basis B von V liefert einen linearen Isomorphismus, $h \leftrightarrow [h]_B$, zwischen dem Vektorraum der Sesquilinearformen auf V und $M_{n \times n}(\mathbb{C})$. Insbesondere ist h durch die Matrix $[h]_B$ eindeutig bestimmt,

$$h(v, w) = [v]_B^*[h]_B[w]_B, \quad v, w \in V,$$

und zu jedem $A \in M_{n \times n}(\mathbb{C})$ existiert eine eindeutig bestimmte Sesquilinearform h auf V mit, sodass $[h]_B = A$. Darüber hinaus ist h genau dann Hermitesch, wenn $[h]_B^* = [h]_B$. Ist C eine weitere geordnete Basis von V , dann gilt

$$[\beta]_B = T_{CB}^* [\beta]_C T_{CB},$$

wobei $T_{CB} \in \text{GL}_n(\mathbb{C})$ die Matrix zum Basiswechsel von B nach C bezeichnet.

BEWEIS. Analog zum Beweis von Proposition VII.1.9. \square

VII.1.51. BEMERKUNG (Unitäre Gruppe). Ist $h: V \times V \rightarrow \mathbb{C}$ eine Hermitesche Form, dann bildet

$$U(V, h) := \{ \varphi \in \text{GL}(V) \mid \forall v, w \in V : h(\varphi(v), \varphi(w)) = h(v, w) \}$$

eine Gruppe. Ist V endlich dimensional, B eine Basis von V und $A := [h]_B$, dann schränkt sich der Gruppenisomorphismus $\text{GL}(V) \cong \text{GL}_n(\mathbb{C})$, $\varphi \leftrightarrow [\varphi]_{BB}$, siehe Bemerkung IV.6.17, zu einem Gruppenisomorphismus,

$$U(V, h) \cong \{ X \in \text{GL}_n(\mathbb{C}) \mid X^* A X = A \},$$

ein. Dies folgt aus Proposition VII.1.50, denn

$$\begin{aligned} h(\varphi(v), \varphi(w)) &= [\varphi(v)]_B^* [h]_B [\varphi(w)]_B \\ &= ([\varphi]_{BB} [v]_B)^* [h]_B ([\varphi]_{BB} [w]_B) = [v]_B^* \left([\varphi]_{BB}^* [h]_B [\varphi]_{BB} \right) [w]_B \end{aligned}$$

und $h(v, w) = [v]_B^* [h]_B [w]$.

VII.1.52. DEFINITION (Quadratische Form). Ist $h: V \times V \rightarrow \mathbb{C}$ eine Hermitesche Form, dann wird $q: V \rightarrow \mathbb{R} \subseteq \mathbb{C}$, $q(v) := h(v, v)$, als die mit h assoziierte quadratische Form bezeichnet. Beachte, dass $h(v, v)$ reell ist, denn $h(v, v) = \overline{h(v, v)}$. Auch ist q homogen vom Grad zwei, d.h.

$$q(\lambda v) = |\lambda|^2 q(v), \quad \lambda \in \mathbb{C}, \quad v \in V.$$

VII.1.53. PROPOSITION (Polarisierungsidentität). Sei $h: V \times V \rightarrow \mathbb{C}$ eine Hermitesche Form und $q(v) = h(v, v)$ die damit assoziierte quadratische Form. Dann gilt

$$h(v, w) = \frac{1}{4} (q(v+w) - q(v-w)) - \mathbf{i} \frac{1}{4} (q(v+iw) - q(v-iw)).$$

Insbesondere kann die Hermitesche Form h aus der quadratischen Form q zurückgewonnen werden.

BEWEIS. Durch direktes Nachrechnen, siehe Aufgabe 95. \square

VII.1.54. BEMERKUNG. Aus der Polarisierungsidentität erhalten wir

$$U(V, h) = \{ \varphi \in \text{GL}(V) \mid \forall v \in V : q(\varphi(v)) = q(v) \},$$

wobei $q(v) = h(v, v)$ die mit h assoziierte quadratische Form bezeichnet.

VII.1.55. DEFINITION (Definitheit). Eine Hermitesche Form auf einem komplexen Vektorraum, $h: V \times V \rightarrow \mathbb{C}$, und die assoziierte quadratische Form, $q: V \rightarrow \mathbb{R}$, $q(v) = h(v, v)$, heißen:

- (a) *positiv semidefinit*, in Zeichen $h \geq 0$, falls $h(v, v) \geq 0$, für alle $v \in V$.
- (b) *positiv definit*, in Zeichen $h > 0$, falls $h(v, v) > 0$, für alle $0 \neq v \in V$.
- (c) *negativ semidefinit*, in Zeichen $h \leq 0$, falls $h(v, v) \leq 0$, für alle $v \in V$.
- (d) *negativ definit*, in Zeichen $h < 0$, falls $h(v, v) < 0$, für alle $0 \neq v \in V$.
- (e) *indefinit*, falls es $v, w \in V$ gibt, sodass $h(v, v) > 0$ und $h(w, w) < 0$.

Eine Hermitesche Matrix $A = A^*$ wird positiv bzw. negativ (semi)definit genannt, falls die assoziierte Hermitesche Form, $h_A: \mathbb{C}^n \times \mathbb{C}^n \rightarrow \mathbb{R}$, $h_A(x, y) = x^*Ay$, positiv bzw. negativ (semi)definit ist, d.h. $A \geq 0$ falls $x^*Ax \geq 0$ für alle $x \in \mathbb{C}^n$; $A > 0$ falls $x^*Ax > 0$ für alle $0 \neq x \in \mathbb{C}^n$; $A \leq 0$ falls $x^*Ax \leq 0$ für alle $x \in \mathbb{C}^n$; und $A < 0$ falls $x^*Ax < 0$ für alle $0 \neq x \in \mathbb{C}^n$.

VII.1.56. BEISPIEL. Das standard innere Produkt auf \mathbb{C}^n ,

$$\langle x, y \rangle = \bar{x}_1y_1 + \cdots + \bar{x}_ny_n = x^*y = x^* \begin{pmatrix} 1 & & \\ & \ddots & \\ & & 1 \end{pmatrix} y, \quad x, y \in \mathbb{C}^n,$$

ist eine positiv definite Hermitesche Form auf \mathbb{C}^n , denn

$$\langle x, x \rangle = \sum_{i=1}^n \bar{x}_ix_i = \sum_{i=1}^n |x_i|^2 > 0,$$

für jedes $0 \neq x \in \mathbb{C}^n$.

VII.1.57. BEISPIEL. Die Abbildung

$$\langle -, - \rangle: M_{m \times n}(\mathbb{C}) \times M_{m \times n}(\mathbb{C}) \rightarrow \mathbb{C}, \quad \langle A, B \rangle := \text{tr}(A^*B)$$

ist eine positiv definite Hermitesche Form auf $M_{m \times n}(\mathbb{C})$, denn

$$\text{tr}(A^*A) = \sum_{j=1}^n \sum_{i=1}^m \bar{A}_{ij}A_{ij} = \sum_{j=1}^n \sum_{i=1}^m |A_{ij}|^2 > 0$$

für jedes $0 \neq A \in M_{m \times n}(\mathbb{C})$.

VII.1.58. BEISPIEL. Ist $I \subseteq \mathbb{R}$ ein kompaktes Intervall und bezeichnet $C^0(I, \mathbb{C})$ den Vektorraum der stetigen, komplexwertigen Funktionen auf I , dann ist

$$\langle -, - \rangle: C^0(I, \mathbb{C}) \times C^0(I, \mathbb{C}) \rightarrow \mathbb{C}, \quad \langle f, g \rangle := \int_I \overline{f(t)}g(t)dt,$$

eine positiv definite Hermitesche Form auf $C^0(I, \mathbb{C})$, denn

$$\langle f, f \rangle = \int_I \overline{f(t)}f(t) dt = \int_I |f(t)|^2 dt > 0,$$

für jedes $0 \neq f \in C^0(I, \mathbb{C})$.

VII.1.59. SATZ (Trägheitssatz von Sylvester). Sei V ein n -dimensionaler komplexer Vektorraum und $h: V \times V \rightarrow \mathbb{R}$ eine Hermitesche Form. Dann existiert eine geordnete Basis B von V , sodass

$$[h]_B = \begin{pmatrix} I_p & & \\ & -I_q & \\ & & 0 \end{pmatrix}.$$

Für jede solche Basis gilt

$$p = \max\{\dim(W) \mid W \text{ Teilraum von } V \text{ mit } h|_W > 0\},$$

$$q = \max\{\dim(W) \mid W \text{ Teilraum von } V \text{ mit } h|_W < 0\}.$$

Insbesondere sind die Zahlen p und q unabhängig von der Basis B . Das Paar (p, q) wird als Signatur der Hermiteschen Form h bezeichnet.

Eine Hermitesche Form auf einem n -dimensionalen komplexen Vektorraum ist genau dann positiv definit, wenn sie Signatur $(n, 0)$ hat. Sie ist genau dann negativ definit, wenn sie Signatur $(0, n)$ hat. Die Menge der positiv definiten Hermiteschen Formen ist konvex, und auch die Menge der negativ definiten Hermiteschen Formen ist konvex.

VII.1.60. KOROLLAR. Zwei Hermitesche $(n \times n)$ -Matrizen A und B haben genau dann gleiche Signatur, wenn eine invertierbare Matrix $S \in \text{GL}_n(\mathbb{C})$ existiert, sodass $B = S^*AS$.

VII.1.61. KOROLLAR. Sei $h: V \times V \rightarrow \mathbb{C}$ eine positiv definite Hermitesche Form auf einem n -dimensionalen komplexen Vektorraum V . Dann existiert eine Basis B von V , sodass $[h]_B = I_n$, d.h.

$$h(v, w) = [v]_B^*[w]_B, \quad v, w \in V.$$

Jede solche Basis liefert einen Gruppenisomorphismus,

$$U(V, \beta) \cong U_n := \{A \in M_{n \times n}(\mathbb{C}) \mid A^*A = I_n\}, \quad \varphi \leftrightarrow [\varphi]_{BB}.$$

VII.1.62. SATZ (Sylvester Kriterium). Sei V ein endlich dimensionaler komplexer Vektorraum. Eine Hermitesche Form $h: V \times V \rightarrow \mathbb{R}$ ist genau dann positiv definit, wenn alle Hauptminoren der Matrix $[h]_B$ positiv sind, bezüglich einer (und dann jeder) Basis B von V .

BEWEIS. Analog zum Beweis von Satz VII.1.42. □

VII.1.63. BEISPIEL. Wir wollen die Signatur der mit

$$A = \begin{pmatrix} 1 & 1 + \mathbf{i} \\ 1 - \mathbf{i} & 1 \end{pmatrix}$$

assozierten Hermiteschen Form $h_A(v, w) = v^*Aw$ auf \mathbb{C}^2 bestimmen. Durch Erganzen auf vollstandige Quadrate erhalten wir

$$\begin{aligned} q \begin{pmatrix} x \\ y \end{pmatrix} &= \begin{pmatrix} x \\ y \end{pmatrix}^* \begin{pmatrix} 1 & 1 + \mathbf{i} \\ 1 - \mathbf{i} & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \bar{x}x + \bar{y}y + (1 + \mathbf{i})\bar{x}y + (1 - \mathbf{i})\bar{y}x \\ &= \overline{(x + (1 + \mathbf{i})y)}(x + (1 + \mathbf{i})y) - \bar{y}y, \end{aligned}$$

die Hermitesche Form h_A hat daher Signatur $(1, 1)$. Aus obiger Rechnung folgt

$$v^*Av = q(v) = (Tv)^* \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} (Tv) \quad \text{mit} \quad T = \begin{pmatrix} 1 & 1 + \mathbf{i} \\ 0 & 1 \end{pmatrix},$$

also $A = T^* \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} T$, siehe Proposition VII.1.53, und daher

$$S^*AS = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad \text{wobei} \quad S = T^{-1} = \begin{pmatrix} 1 & -1 - \mathbf{i} \\ 0 & 1 \end{pmatrix}.$$

VII.2. Innere Produkte.

VII.2.1. DEFINITION (Euklidische Vektorrume). Unter einem *inneren Produkt* auf einem reellen Vektorraum V verstehen wir eine positiv definite symmetrische Bilineaform auf V , d.h. eine Abbildung $\langle -, - \rangle: V \times V \rightarrow \mathbb{R}$, mit folgenden Eigenschaften:

- (a) $\langle v + v', w \rangle = \langle v, w \rangle + \langle v', w \rangle$, $\langle \lambda v, w \rangle = \lambda \langle v, w \rangle$, fur alle $v, v', w \in V$, $\lambda \in \mathbb{R}$.
- (b) $\langle v, w + w' \rangle = \langle v, w \rangle + \langle v, w' \rangle$, $\langle v, \lambda w \rangle = \lambda \langle v, w \rangle$, fur alle $v, w, w' \in V$, $\lambda \in \mathbb{R}$.
- (c) $\langle w, v \rangle = \langle v, w \rangle$, fur alle $v, w \in V$.
- (d) $\langle v, v \rangle > 0$, fur alle $0 \neq v \in V$.

Ein reeller Vektorraum, der mit einem inneren Produkt ausgestattet ist, wird *Euklidischer Vektorraum* genannt.⁸

VII.2.2. DEFINITION (Unitare Vektorrume). Unter einem *inneren Produkt* auf einem komplexen Vektorraum V verstehen wir eine positiv definite Hermitesche Form auf V , d.h. eine Abbildung $\langle -, - \rangle: V \times V \rightarrow \mathbb{C}$, mit folgenden Eigenschaften:

- (a) $\langle v + v', w \rangle = \langle v, w \rangle + \langle v', w \rangle$, $\langle \lambda v, w \rangle = \bar{\lambda} \langle v, w \rangle$, fur alle $v, v', w \in V$, $\lambda \in \mathbb{C}$.
- (b) $\langle v, w + w' \rangle = \langle v, w \rangle + \langle v, w' \rangle$, $\langle v, \lambda w \rangle = \lambda \langle v, w \rangle$, fur alle $v, w, w' \in V$, $\lambda \in \mathbb{C}$.
- (c) $\langle w, v \rangle = \overline{\langle v, w \rangle}$, fur alle $v, w \in V$.
- (d) $\langle v, v \rangle > 0$, fur alle $0 \neq v \in V$.

Ein komplexer Vektorraum, der mit einem inneren Produkt ausgestattet ist, wird *unitarer Vektorraum* genannt.⁹

Ist $\langle -, - \rangle$ ein inneres Produkt auf einem reellen oder komplexen Vektorraum, so schreiben wir

$$\|v\| := \sqrt{\langle v, v \rangle}, \quad v \in V.$$

⁸Diese Forderungen sind redundant, etwa lasst sich (b) sofort aus (a) und (c) folgern.

⁹Auch hier folgt (b) aus (a) und (c).

Beachte, dass die Wurzel wohldefiniert ist, da $\langle v, v \rangle \geq 0$. Beachte auch, dass $\|v\|^2 = \langle v, v \rangle = q(v)$ gerade die mit $\langle -, - \rangle$ assoziierte quadratische Form ist. Aus den Propositionen VII.1.22 und VII.1.53 erhalten wir daher:

VII.2.3. PROPOSITION (Polarisierungsidentitäten).

a) Für je zwei Vektoren v und w eines Euklidischen Vektorraums gilt:

$$\langle v, w \rangle = \frac{1}{2}(\|v + w\|^2 - \|v\|^2 - \|w\|^2) = \frac{1}{4}(\|v + w\|^2 + \|v - w\|^2).$$

b) Für je zwei Vektoren v und w eines unitären Vektorraums gilt:

$$\langle v, w \rangle = \frac{1}{4}(\|(v + w)\|^2 - \|v - w\|^2 - \|\mathbf{i}v + \mathbf{i}w\|^2 + \|\mathbf{i}v - \mathbf{i}w\|^2).$$

VII.2.4. BEISPIEL. Der Vektorraum \mathbb{R}^n mit dem standard inneren Produkt,

$$\langle -, - \rangle: \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}, \quad \langle x, y \rangle = x_1y_1 + \cdots + x_ny_n,$$

ist ein Euklidischer Vektorraum mit Norm

$$\|x\|^2 = |x_1|^2 + \cdots + |x_n|^2$$

VII.2.5. BEISPIEL. Der Vektorraum \mathbb{C}^n mit dem standard inneren Produkt,

$$\langle -, - \rangle: \mathbb{C}^n \times \mathbb{C}^n \rightarrow \mathbb{C}, \quad \langle x, y \rangle = \bar{x}_1y_1 + \cdots + \bar{x}_ny_n,$$

ist ein unitärer Vektorraum mit Norm

$$\|x\|^2 = |x_1|^2 + \cdots + |x_n|^2.$$

VII.2.6. BEISPIEL. Der Vektorraum der stetigen Funktionen auf einem kompakten Intervall, $C^0(I, \mathbb{C})$, wird durch das innere Produkt,

$$\langle f, g \rangle = \int_I \overline{f(t)}g(t)dt,$$

zu einem unitären Vektorraum mit Norm

$$\|f\|^2 = \int_I |f(t)|^2 dt.$$

VII.2.7. BEISPIEL. Der Vektorraum der *quadratsummierbaren* Folgen,

$$l^2 := \{x: \mathbb{N} \rightarrow \mathbb{C} \mid \sum_{n \in \mathbb{N}} |x_n|^2 < \infty\},$$

ein Teilraum des Vektorraums aller Folgen, wird durch das innere Produkt

$$\langle x, y \rangle := \sum_{n \in \mathbb{N}} \bar{x}_ny_n, \quad x, y \in l^2(\mathbb{N}), \quad (\text{VII.8})$$

zu einem unitären Vektorraum mit Norm

$$\|x\|^2 = \sum_{n \in \mathbb{N}} |x_n|^2.$$

Mit Hilfe der Cauchy–Schwarz Ungleichung, siehe Proposition VII.2.9 unten, lässt sich zeigen, dass die Reihe (VII.8) absolut konvergent ist.

VII.2.8. PROPOSITION (Satz von Pythagoras). *Sei V ein Euklidischer oder unitärer Vektorraum. Sind $v, w \in V$ mit $\langle v, w \rangle = 0$ so gilt*

$$\|v + w\|^2 = \|v\|^2 + \|w\|^2.$$

BEWEIS. Aus der Bilinearität des inneren Produkts folgt sofort

$$\|v + w\|^2 = \langle v + w, v + w \rangle = \langle v, v \rangle + \langle v, w \rangle + \langle w, v \rangle + \langle w, w \rangle, = \|v\|^2 + \|w\|^2,$$

denn nach Voraussetzung ist $\langle v, w \rangle = 0$ und daher auch $\langle w, v \rangle = \overline{\langle v, w \rangle} = 0$. \square

VII.2.9. PROPOSITION (Cauchy–Schwarz Ungleichung). *Sei V ein Euklidischer oder unitärer Vektorraum mit innerem Produkt $\langle -, - \rangle$. Dann gilt*

$$|\langle v, w \rangle| \leq \|v\| \|w\|, \quad (\text{VII.9})$$

für alle $v, w \in V$. Gleichheit tritt in (VII.9) genau dann ein, wenn v und w linear abhängig sind.

BEWEIS. O.B.d.A. sei $v \neq 0$. Setzen wir $\tilde{v} := \frac{\langle v, w \rangle}{\|v\|^2} v$ und $\tilde{w} = w - \tilde{v}$, dann gilt

$$w = \tilde{v} + \tilde{w}, \quad \langle \tilde{v}, \tilde{w} \rangle = 0, \quad \text{und} \quad \|\tilde{v}\|^2 = \frac{|\langle v, w \rangle|^2}{\|v\|^2}. \quad (\text{VII.10})$$

Die erste Gleichung ist trivial, die dritte folgt aus

$$\|\tilde{v}\|^2 = \langle \tilde{v}, \tilde{v} \rangle = \left\langle \frac{\langle v, w \rangle}{\|v\|^2} v, \frac{\langle v, w \rangle}{\|v\|^2} v \right\rangle = \frac{\overline{\langle v, w \rangle} \langle v, w \rangle}{\|v\|^2 \|v\|^2} \langle v, v \rangle = \frac{|\langle v, w \rangle|^2}{\|v\|^2},$$

und die zweite via

$$\begin{aligned} \langle \tilde{v}, \tilde{w} \rangle &= \langle \tilde{v}, w \rangle - \langle \tilde{v}, \tilde{v} \rangle = \left\langle \frac{\langle v, w \rangle}{\|v\|^2} v, w \right\rangle - \|\tilde{v}\|^2 \\ &= \frac{\overline{\langle v, w \rangle}}{\|v\|^2} \langle v, w \rangle - \|\tilde{v}\|^2 = \frac{|\langle v, w \rangle|^2}{\|v\|^2} - \|\tilde{v}\|^2 = 0. \end{aligned}$$

Aus den Relationen in (VII.10) und Proposition VII.2.8 erhalten wir

$$\|w\|^2 = \|\tilde{v} + \tilde{w}\|^2 = \|\tilde{v}\|^2 + \|\tilde{w}\|^2 \geq \|\tilde{v}\|^2 = \frac{|\langle v, w \rangle|^2}{\|v\|^2}$$

also (VII.9). Gleichheit kann nur eintreten, wenn $\tilde{w} = 0$ gilt, d.h. wenn w ein Vielfaches von v ist. \square

VII.2.10. PROPOSITION (Dreiecksungleichung). *Sei V ein Euklidischer oder unitärer Vektorraum. Dann gilt für alle $v, w \in V$,*

$$\|v + w\| \leq \|v\| + \|w\|.$$

BEWEIS. Aus der Bilinearität des inneren Produkts und der Cauchy–Schwarz Ungleichung in Proposition VII.2.9 erhalten wir:

$$\begin{aligned}
\|v + w\|^2 &= \langle v + w, v + w \rangle \\
&= \langle v, v \rangle + \langle v, w \rangle + \langle w, v \rangle + \langle w, w \rangle \\
&= \|v\|^2 + \langle v, w \rangle + \overline{\langle v, w \rangle} + \|w\|^2 \\
&= \|v\|^2 + 2 \operatorname{Re} \langle v, w \rangle + \|w\|^2 \\
&\leq \|v\|^2 + 2|\langle v, w \rangle| + \|w\|^2 \\
&\leq \|v\|^2 + 2\|v\|\|w\| + \|w\|^2 \\
&= (\|v\| + \|w\|)^2.
\end{aligned}$$

Mit der Monotonie der Wurzel folgt daher $\|v + w\| \leq \|v\| + \|w\|$. \square

VII.2.11. DEFINITION (Norm). Sei V ein reeller oder komplexer Vektorraum. Unter einer Norm auf V verstehen wir eine Abbildung $\|-\|: V \rightarrow \mathbb{R}$ mit folgenden Eigenschaften:

- (a) $\|v + w\| \leq \|v\| + \|w\|$, für alle $v, w \in V$. (Dreiecksungleichung)
- (b) $\|\lambda v\| = |\lambda| \|v\|$, für alle $v \in V$ und $\lambda \in \mathbb{R}$ bzw. $\lambda \in \mathbb{C}$.
- (c) $\forall v \in V : \|v\| = 0 \Rightarrow v = 0$.¹⁰

Unter einem normierten Vektorraum verstehen wir einen Vektorraum, der mit einer Norm ausgestattet ist. Beachte, $\|v\| \geq 0$ für alle $v \in V$, denn aus (a) und (b) folgt $0 = \|0\| = \|v - v\| \leq \|v\| + \|(-1)v\| = 2\|v\|$.

VII.2.12. PROPOSITION. Sei V ein Euklidischer oder unitärer Vektorraum mit innerem Produkt $\langle -, - \rangle$. Dann ist $\|v\| = \sqrt{\langle v, v \rangle}$ eine Norm auf V , für die die Parallelogrammgleichung gilt, d.h. für alle $v, w \in V$ haben wir

$$\|v + w\|^2 + \|v - w\|^2 = 2(\|v\|^2 + \|w\|^2).$$

BEWEIS. Die Dreiecksungleichung haben wir in Proposition VII.2.10 gezeigt. Weiters gilt

$$\|\lambda v\|^2 = \langle \lambda v, \lambda v \rangle = \bar{\lambda} \lambda \langle v, v \rangle = |\lambda|^2 \|v\|^2 = (|\lambda| \|v\|)^2,$$

also auch $\|\lambda v\| = |\lambda| \|v\|$. Ist $\|v\| = 0$, dann gilt $\langle v, v \rangle = 0$, also $v = 0$ wegen der positiv Definitheit des inneren Produkts. Aus der Bilinearität des inneren Produkts erhalten wir:

$$\begin{aligned}
\|v + w\|^2 &= \langle v + w, v + w \rangle = \|v\|^2 + \|w\|^2 + \langle v, w \rangle + \langle w, v \rangle \\
\|v - w\|^2 &= \langle v - w, v - w \rangle = \|v\|^2 + \|w\|^2 - \langle v, w \rangle - \langle w, v \rangle
\end{aligned}$$

Aufaddieren liefert die Parallelogrammgleichung. \square

VII.2.13. DEFINITION (Metrische Räume). Unter einer Metrik auf einer Menge X verstehen wir eine Abbildung $d: X \times X \rightarrow \mathbb{R}$ mit folgenden Eigenschaften:

¹⁰Aus (b) folgt sofort $\|0\| = 0$, es gilt daher $\forall v \in V : \|v\| = 0 \Leftrightarrow v = 0$.

- (a) $d(x, z) \leq d(x, y) + d(y, z)$, für alle $x, y, z \in X$. (Dreiecksungleichung)
 (b) $d(x, y) = d(y, x)$, für alle $x, y \in X$. (Symmetrie)
 (c) $d(x, y) = 0 \Leftrightarrow x = y$.

Unter einem *metrischen Raum* verstehen wir eine Menge X , die mit einer Metrik ausgestattet ist. Beachte $d(x, y) \geq 0$ für alle $x, y \in X$, denn aus (a), (b) und (c) folgt $0 = d(x, x) \leq d(x, y) + d(y, x) = 2d(x, y)$.

VII.2.14. PROPOSITION. *Ist V ein normierter Raum, dann definiert $d(v, w) := \|w - v\|$ eine Metrik auf V .*

BEWEIS. Die Dreiecksungleichung folgt aus Definition VII.2.11(a),

$$d(u, w) = \|w - u\| = \|w - v + v - u\| \leq \|w - v\| + \|v - u\| = d(v, w) + d(u, v).$$

Symmetrie folgt aus Definition VII.2.11(b),

$$d(v, w) = \|w - v\| = \|(-1)(v - w)\| = |(-1)|\|v - w\| = \|v - w\| = d(w, v).$$

Schließlich gilt $d(v, w) = 0$ genau dann wenn $\|w - v\| = 0$, d.h. genau dann wenn $w - v = 0$, siehe Definition VII.2.11(c). \square

VII.2.15. BEMERKUNG. Nicht jede Norm kommt von einem inneren Produkt. Etwa definiert

$$\|x\|_p := (|x_1|^p + \dots + |x_n|^p)^{1/p}$$

für jedes $1 \leq p < \infty$ eine Norm auf \mathbb{R}^n bzw. \mathbb{C}^n , die aber nur im Fall $p = 2$ der Parallelogrammgleichung genügt. Ist I ein kompaktes Intervall, so ist

$$\|f\|_p := \left(\int_I |f(t)|^p dt \right)^{1/p}$$

eine Norm auf dem Vektorraum der stetigen Funktionen, $C^0(I, \mathbb{C})$ bzw. $C^0(I, \mathbb{R})$, die im Fall $p \neq 2$ nicht der Parallelogrammgleichung genügt. Auch die Supremumsnormen

$$\|x\|_\infty := \max\{|x_1|, \dots, |x_n|\} \quad \text{bzw.} \quad \|f\|_\infty := \max_{t \in I} |f(t)|$$

auf \mathbb{C}^n bzw. $C^0(I, \mathbb{C})$ sind Normen die nicht der Parallelogrammgleichung genügen und daher nicht von einem inneren Produkt stammen. Ein weiteres wichtiges Beispiel ist die sogenannte *Operatornorm* linearer Abbildungen $\varphi: V \rightarrow W$ zwischen Euklidischen oder unitären Vektorräumen,

$$\|\varphi\| := \sup_{0 \neq v \in V} \frac{\|\varphi(v)\|}{\|v\|} = \sup_{v \in V, \|v\|=1} \|\varphi(v)\|.$$

Es lässt sich zeigen, dass im Fall $\dim(V) < \infty$ dieses Supremum endlich ist und eine Norm auf $L(V, W)$ definiert, die i.A. nicht die Parallelogrammgleichung erfüllt.

VII.2.16. DEFINITION (Winkel). Ist V ein Euklidischer Vektorraum und sind $v, w \in V$, $v \neq 0 \neq w$, dann gilt nach Proposition VII.2.9 $-1 \leq \frac{\langle v, w \rangle}{\|v\|\|w\|} \leq 1$, es existiert daher ein eindeutiges $\alpha \in [0, \pi]$, sodass¹¹

$$\cos(\alpha) = \frac{\langle v, w \rangle}{\|v\|\|w\|}.$$

Diese Zahl α wird als *Winkel* zwischen v und w bezeichnet.

VII.2.17. PROPOSITION. Seien $v \neq 0 \neq w$ zwei Vektoren eines Euklidischen Vektorraums und $\alpha \in [0, \pi]$ der Winkel zwischen v und w . Dann gilt:

- (a) $\alpha = 0$, genau dann wenn $\lambda > 0$ existiert, sodass $w = \lambda v$.
- (b) $\alpha = \pi/2$, genau dann wenn $\langle v, w \rangle = 0$.
- (c) $\alpha = \pi$, genau dann wenn $\lambda < 0$ existiert, sodass $w = \lambda v$.

BEWEIS. Ist $w = \lambda v$ mit $\lambda \in \mathbb{R}$, dann folgt $\lambda \neq 0$ und

$$\cos(\alpha) = \frac{\langle v, w \rangle}{\|v\|\|w\|} = \frac{\langle v, \lambda v \rangle}{\|v\|\|\lambda v\|} = \frac{\lambda \langle v, v \rangle}{|\lambda| \|v\| \|v\|} = \frac{\lambda}{|\lambda|}.$$

Im Fall $\lambda > 0$ erhalten wir $\cos(\alpha) = 1$ und daher $\alpha = 0$. Im Fall $\lambda < 0$ erhalten wir $\cos(\alpha) = -1$, also $\alpha = \pi$. Seien nun $v \neq 0 \neq w$ beliebig und $\alpha = 0$ oder $\alpha = \pi$, d.h. $\cos(\alpha) = \pm 1$. Dann gilt $|\langle v, w \rangle| = \|v\|\|w\|$, nach Proposition VII.2.9 existiert daher $\lambda \in \mathbb{R}$ mit $w = \lambda v$. Dies zeigt (a) und (c). Behauptung (b) ist trivial, $\cos(\pi/2) = 0$. \square

VII.2.18. BEMERKUNG (Cosinussatz). Sind $v \neq 0 \neq w$ zwei Vektoren in einem Euklidischen Vektorraum und bezeichnet α den Winkel zwischen v und w , so gilt

$$\|v - w\|^2 = \|v\|^2 + \|w\|^2 - 2 \cos(\alpha) \|v\| \|w\|,$$

siehe Aufgabe 101.

VII.2.19. DEFINITION (Orthogonalität). Sei V ein Euklidischer oder unitärer Vektorraum. Zwei Vektoren $v, w \in V$ werden *orthogonal* genannt, falls $\langle v, w \rangle = 0$. Wir schreiben in diesem Fall $v \perp w$. Ein Vektor $v \in V$ wird *normiert* genannt, falls $\|v\| = 1$. Unter einem *Orthonormalsystem* in V verstehen wir ein System von Vektoren $v_i \in V$, $i \in I$, sodass $\langle v_i, v_j \rangle = \delta_{ij}$, für alle $i, j \in I$. Eine *Orthonormalbasis* von V ist ein Orthonormalsystem, das eine Basis von V bildet.

VII.2.20. BEISPIEL. Die Standardbasen von \mathbb{R}^n und \mathbb{C}^n sind Orthonormalbasen bezüglich der standard inneren Produkte, siehe Beispiel VII.2.4 und VII.2.5. Für jedes $\theta \in \mathbb{R}$ bilden die beiden Vektoren,

$$b_1 = \begin{pmatrix} \cos \theta \\ \sin \theta \end{pmatrix} \quad \text{und} \quad b_2 = \begin{pmatrix} -\sin \theta \\ \cos \theta \end{pmatrix},$$

eine Orthonormalbasis von \mathbb{R}^2 .

¹¹Der Cosinus schränkt sich zu einer Bijektion $\cos: [0, \pi] \rightarrow [-1, 1]$ ein.

Eine Basis, B , eines endlich dimensionalen Euklidischen oder unitären Vektorraums, V , ist also genau dann eine Orthonormalbasis, wenn die Matrixdarstellung des inneren Produkts bezüglich B durch die Einheitsmatrix gegeben ist, d.h.

$$\langle v, w \rangle = [v]_B^* [w]_B, \quad \text{bzw.} \quad \|v\|^2 = [v]_B^* [v]_B, \quad v, w \in V. \quad (\text{VII.11})$$

Nach den Ergebnissen des vorangehenden Abschnitts besitzt jeder endlich dimensionale Euklidische (Korollar VII.1.39) oder unitäre (Korollar VII.1.26) Vektorraum daher eine Orthonormalbasis.

Ein System von Vektoren b_1, \dots, b_n in \mathbb{R}^n ist genau dann eine Orthonormalbasis, wenn die Matrix $A = (b_1 | \dots | b_n)$ orthogonal ist, d.h.

$$A \in O_n = \{A \in M_{n \times n}(\mathbb{R}) : A^t A = I_n\},$$

denn $(A^t A)_{ij} = b_i^t b_j = \langle b_i, b_j \rangle$. Etwa ist

$$\begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix},$$

für jedes $\theta \in \mathbb{R}$, eine orthogonale Matrix. Beachte, dass jedes $A \in O_n$ invertierbar ist, $A^{-1} = A^t$, es gilt daher auch $AA^t = I_n$. Im vorangehenden Abschnitt haben wir gesehen, dass O_n eine Untergruppe von $\text{GL}_n(\mathbb{R})$ bildet, sie wird die *orthogonale Gruppe* genannt, ihre Elemente werden als *orthogonale Matrizen* bezeichnet. Es gilt $\det(A) = \pm 1$, für jede orthogonale Matrix $A \in O_n$, vgl. Aufgabe 109. Die Determinante liefert einen surjektiven Gruppenhomomorphismus, $\det : O_n \rightarrow \{\pm 1\}$, wobei $\{\pm 1\}$ bezüglich Multiplikation als (abelsche) Gruppe aufgefasst wird. Sein Kern,

$$\text{SO}_n := \{A \in O_n : \det(A) = 1\},$$

bildet eine Untergruppe von O_n , die als *spezielle orthogonale Gruppe* bezeichnet wird. Es lässt sich zeigen, dass O_n und SO_n kompakte Teilmengen von $M_{n \times n}(\mathbb{R})$ sind, vgl. Aufgabe 108.

Analog ist ein System von Vektoren b_1, \dots, b_n in \mathbb{C}^n genau dann eine Orthonormalbasis, wenn die Matrix $A = (b_1 | \dots | b_n)$ unitär ist, d.h.

$$A \in U_n = \{A \in M_{n \times n}(\mathbb{C}) : A^* A = I_n\},$$

denn $(A^* A)_{ij} = b_i^* b_j = \langle b_i, b_j \rangle$. Jedes $A \in U_n$ ist invertierbar, $A^{-1} = A^*$, es gilt daher auch $AA^* = I_n$. Im vorangehenden Abschnitt haben wir gesehen, dass U_n eine Untergruppe von $\text{GL}_n(\mathbb{C})$ bildet, sie wird die *unitäre Gruppe* genannt, ihre Elemente werden als *unitäre Matrizen* bezeichnet. Es gilt $|\det(A)| = 1$, für jede unitäre Matrix $A \in U_n$. Die Determinante liefert einen surjektiven Gruppenhomomorphismus, $\det : U_n \rightarrow S^1 := \{z \in \mathbb{C} : |z| = 1\}$, wobei S^1 bezüglich Multiplikation als (abelsche) Gruppe aufgefasst wird. Sein Kern,

$$\text{SU}_n := \{A \in U_n : \det(A) = 1\},$$

bildet eine Untergruppe von U_n , die als *spezielle unitäre Gruppe* bezeichnet wird. Es lässt sich zeigen, dass U_n und SU_n kompakte Teilmengen von $M_{n \times n}(\mathbb{C}) = \mathbb{C}^{n^2} = \mathbb{R}^{4n^2}$ sind.

VII.2.21. LEMMA. *Jedes Orthonormalsystem eines Euklidischen oder unitären Vektorraums ist linear unabhängig.*

BEWEIS. O.B.d.A. genügt es endliche Orthonormalsysteme zu betrachten. Sei also v_1, \dots, v_n , ein Orthonormalsystem in V . Weiters seien λ_i Skalare, sodass $\lambda_1 v_1 + \dots + \lambda_n v_n = 0$. Für jedes $i = 1, \dots, n$, erhalten wir

$$0 = \langle v_i, 0 \rangle = \langle v_i, \lambda_1 v_1 + \dots + \lambda_n v_n \rangle = \lambda_1 \langle v_i, v_1 \rangle + \dots + \lambda_n \langle v_i, v_n \rangle = \lambda_i,$$

denn $\langle v_i, v_j \rangle = \delta_{ij}$, für alle $1 \leq i, j \leq n$. Dies zeigt, dass v_1, \dots, v_n linear unabhängig sind. \square

VII.2.22. PROPOSITION. *Sei $B = (b_1, \dots, b_n)$ eine Orthonormalbasis eines endlich dimensionalen Euklidischen oder unitären Vektorraums V . Dann gilt*

$$v = \sum_{i=1}^n \langle b_i, v \rangle b_i \quad \text{bzw.} \quad [v]_B = \begin{pmatrix} \langle b_1, v \rangle \\ \vdots \\ \langle b_n, v \rangle \end{pmatrix},$$

und daher auch

$$\|v\|^2 = \sum_{i=1}^n |\langle b_i, v \rangle|^2 \quad \text{und} \quad \langle v, w \rangle = \sum_{i=1}^n \overline{\langle b_i, v \rangle} \langle b_i, w \rangle,$$

für alle $v, w \in V$.

BEWEIS. Da B eine Basis von V bildet, existieren Skalare $\lambda_1, \dots, \lambda_n$, sodass $v = \lambda_1 b_1 + \dots + \lambda_n b_n$. Es folgt

$$\langle b_i, v \rangle = \langle b_i, \sum_{j=1}^n \lambda_j b_j \rangle = \sum_{j=1}^n \lambda_j \langle b_i, b_j \rangle = \sum_{j=1}^n \lambda_j \delta_{ij} = \lambda_i.$$

Die verbleibenden Behauptungen folgen aus (VII.11). \square

VII.2.23. SATZ (Gram–Schmidt Orthonormalisierungsverfahren). *Sei V ein Euklidischer oder unitärer Vektorraum und v_1, \dots, v_n linear unabhängig in V . Dann bilden die rekursiv definierten Vektoren*

$$b_i := \frac{v_i - \sum_{j=1}^{i-1} \langle b_j, v_i \rangle b_j}{\|v_i - \sum_{j=1}^{i-1} \langle b_j, v_i \rangle b_j\|}, \quad 1 \leq i \leq n, \quad (\text{VII.12})$$

ein Orthonormalsystem von V , das den selben Teilraum wie v_1, \dots, v_n aufspannt. Insbesondere ist der Nenner in (VII.12) stets verschieden von Null. Ist v_1, \dots, v_n eine Basis von V , dann bildet b_1, \dots, b_n eine Orthonormalbasis von V .

BEWEIS. Wir führen den Beweis mittels Induktion nach n . Der Induktionsanfang, $n = 1$, ist trivial, denn aus $v_1 \neq 0$ folgt $\|v_1\| \neq 0$, also ist $b_1 = \frac{v_1}{\|v_1\|}$ wohldefiniert, $\|b_1\| = 1$ und $\langle b_1 \rangle = \langle v_1 \rangle$. Für den Induktionsschritt dürfen wir nun annehmen, dass b_1, \dots, b_{n-1} ein wohldefiniertes Orthonormalsystem bildet, für

das $\langle b_1, \dots, b_{n-1} \rangle = \langle v_1, \dots, v_{n-1} \rangle$ gilt. Betrachten wir $\tilde{b}_n := v_n - \sum_{j=1}^{n-1} \langle b_j, v_n \rangle b_j$, dann gilt zunächst

$$\langle b_1, \dots, b_{n-1}, \tilde{b}_n \rangle = \langle b_1, \dots, b_{n-1}, v_n \rangle = \langle v_1, \dots, v_{n-1}, v_n \rangle,$$

wegen der linearen Unabhängigkeit der Vektoren v_1, \dots, v_n muss daher $\tilde{b}_n \neq 0$ sein. Daher ist $b_n = \frac{\tilde{b}_n}{\|\tilde{b}_n\|}$ wohldefiniert und $\langle b_1, \dots, b_n \rangle = \langle v_1, \dots, v_n \rangle$. Nach Konstruktion ist $\|b_n\| = 1$ und

$$\langle b_k, \tilde{b}_n \rangle = \langle b_k, v_n \rangle - \sum_{j=1}^{n-1} \langle b_j, v_n \rangle \underbrace{\langle b_k, b_j \rangle}_{=\delta_{kj}} = \langle b_k, v_n \rangle - \langle b_k, v_n \rangle = 0, \quad 1 \leq k < n,$$

also auch $\langle b_k, b_n \rangle = \frac{1}{\|b_n\|} \langle b_k, \tilde{b}_n \rangle = 0$, für alle $1 \leq k < n$. \square

VII.2.24. KOROLLAR. a) Jeder endlich dimensionale Euklidische oder unitäre Vektorraum besitzt eine Orthonormalbasis.

b) Jedes Orthonormalsystem eines endlich dimensionalen Euklidischen oder unitären Vektorraums kann zu einer Orthonormalbasis erweitert werden.

BEWEIS. Sei also v_1, \dots, v_k ein Orthonormalsystem in V . Wir ergänzen zu einer Basis, $v_1, \dots, v_k, v_{k+1}, \dots, v_n$ von V und wenden darauf das Gram–Schmidt Orthonormalisierungsverfahren an. Dies liefert eine Orthonormalbasis b_1, \dots, b_n von V . Da v_1, \dots, v_k ein Orthonormalsystem bildet gilt $b_i = v_i$ für $i = 1, \dots, k$, vgl. (VII.12). Dies zeigt (b). Behauptung (a) ist ein Spezialfall ($k = 0$), wir haben die Existenz von Orthonormalbasen aber bereits im vorangehenden Abschnitt (mit ähnlichen Methoden) gezeigt, vgl. die Korollare VII.1.39 und VII.1.26. \square

VII.2.25. BEISPIEL. Betrachte den Euklidischen Vektorraum \mathbb{R}^3 mit dem standard inneren Produkt und die von den Vektoren

$$v_1 = \begin{pmatrix} 3 \\ 4 \\ 0 \end{pmatrix}, \quad v_2 = \begin{pmatrix} 6 \\ 8 \\ 7 \end{pmatrix}$$

aufgespannte Ebene $W = \langle v_1, v_2 \rangle$ in \mathbb{R}^3 . Wir wollen eine Orthonormalbasis von W bestimmen. Mit Hilfe des Gram–Schmidt Orthonormalisierungsverfahrens erhalten wir

$$b_1 = \frac{v_1}{\|v_1\|} = \frac{1}{5} \begin{pmatrix} 3 \\ 4 \\ 0 \end{pmatrix}$$

$$\tilde{b}_2 = v_2 - \langle b_1, v_2 \rangle b_1 = \begin{pmatrix} 6 \\ 8 \\ 7 \end{pmatrix} - \frac{50}{25} \begin{pmatrix} 3 \\ 4 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 7 \end{pmatrix}, \quad b_2 = \frac{\tilde{b}_2}{\|\tilde{b}_2\|} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

Nach Satz VII.2.23 bildet daher b_1, b_2 eine Orthonormalbasis von W .

VII.2.26. BEISPIEL. Betrachte den Euklidischen Vektorraum \mathbb{R}^4 mit dem standard inneren Produkt und den von den Vektoren

$$v_1 = \begin{pmatrix} 0 \\ 1 \\ 2 \\ 3 \end{pmatrix}, \quad v_2 = \begin{pmatrix} 1 \\ 2 \\ 3 \\ 0 \end{pmatrix}, \quad v_3 = \begin{pmatrix} 3 \\ 0 \\ 1 \\ 2 \end{pmatrix}$$

aufgespannten Teilraum $W = \langle v_1, v_2, v_3 \rangle$ von \mathbb{R}^4 . Wir wollen eine Orthonormalbasis von W bestimmen. Mit Hilfe des Gram–Schmidt Orthonormalisierungsverfahrens erhalten wir

$$b_1 = \frac{v_1}{\|v_1\|} = \frac{1}{\sqrt{14}} \begin{pmatrix} 0 \\ 1 \\ 2 \\ 3 \end{pmatrix}$$

$$\tilde{b}_2 = v_2 - \langle b_1, v_2 \rangle b_1 = \begin{pmatrix} 1 \\ 2 \\ 3 \\ 0 \end{pmatrix} - \frac{8}{14} \begin{pmatrix} 0 \\ 1 \\ 2 \\ 3 \end{pmatrix} = \frac{1}{7} \begin{pmatrix} 7 \\ 10 \\ 13 \\ -12 \end{pmatrix}, \quad b_2 = \frac{\tilde{b}_2}{\|\tilde{b}_2\|} = \frac{1}{\sqrt{462}} \begin{pmatrix} 7 \\ 10 \\ 13 \\ -12 \end{pmatrix}$$

$$\tilde{b}_3 = v_3 - \langle b_1, v_3 \rangle b_1 - \langle b_2, v_3 \rangle b_2 = \begin{pmatrix} 3 \\ 0 \\ 1 \\ 2 \end{pmatrix} - \frac{8}{14} \begin{pmatrix} 0 \\ 1 \\ 2 \\ 3 \end{pmatrix} - \frac{10}{462} \begin{pmatrix} 7 \\ 10 \\ 13 \\ -12 \end{pmatrix} = \frac{1}{33} \begin{pmatrix} 94 \\ -26 \\ -14 \\ 18 \end{pmatrix}$$

$$b_3 = \frac{\tilde{b}_3}{\|\tilde{b}_3\|} = \frac{1}{4\sqrt{627}} \begin{pmatrix} 94 \\ -26 \\ -14 \\ 18 \end{pmatrix}$$

Nach Satz VII.2.23 ist daher b_1, b_2, b_3 eine Orthonormalbasis von W .

VII.2.27. BEISPIEL. Wir betrachten \mathbb{R}^3 mit dem inneren Produkt

$$\langle x, y \rangle = x^t A y, \quad \text{wobei} \quad A = \begin{pmatrix} 1 & -2 & 1 \\ -2 & 5 & -4 \\ 1 & -4 & 6 \end{pmatrix}. \quad (\text{VII.13})$$

Aus dem Sylvester Kriterium folgt sofort, dass dies tatsächlich positiv definit ist. Wir wollen eine Orthonormalbasis von \mathbb{R}^3 bestimmen und wenden dazu das Gram–Schmidt Orthonormalisierungsverfahren auf die Standardbasis e_1, e_2, e_3 an.

$$b_1 = \frac{e_1}{\|e_1\|} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$$

denn $\|e_1\|^2 = e_1^t A e_1 = 1$.

$$\tilde{b}_2 = e_2 - \langle b_1, e_2 \rangle b_1 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} + 2 \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix}, \quad b_2 = \frac{\tilde{b}_2}{\|\tilde{b}_2\|} = \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix},$$

denn $\langle b_1, e_2 \rangle = b_1^t A e_2 = -2$ und $\|\tilde{b}_2\|^2 = \tilde{b}_2^t A \tilde{b}_2 = 1$.

$$\tilde{b}_3 = e_3 - \langle b_1, e_3 \rangle b_1 - \langle b_2, e_3 \rangle b_2 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} - \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + 2 \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 3 \\ 2 \\ 1 \end{pmatrix},$$

$$b_3 = \frac{\tilde{b}_3}{\|\tilde{b}_3\|} = \begin{pmatrix} 3 \\ 2 \\ 1 \end{pmatrix},$$

denn $\langle b_1, e_3 \rangle = b_1^t A e_3 = 1$, $\langle b_2, e_3 \rangle = b_2^t A e_3 = -2$ und $\|\tilde{b}_3\|^2 = \tilde{b}_3^t A \tilde{b}_3 = 1$. Nach Satz VII.2.23 ist daher b_1, b_2, b_3 eine Orthonormalbasis von \mathbb{R}^3 bezüglich des inneren Produkts (VII.13).

VII.2.28. KOROLLAR (QR-Zerlegung).

a) Jede invertierbare Matrix $A \in \text{GL}_n(\mathbb{R})$ lässt sich auf eindeutige Weise in der Form $A = QR$ schreiben, wobei $Q \in O_n$ und $R \in M_{n \times n}(\mathbb{R})$ eine (invertierbare) obere Dreiecksmatrix mit positiven Diagonaleinträgen ist.

b) Jede invertierbare Matrix $A \in \text{GL}_n(\mathbb{C})$ lässt sich auf eindeutige Weise in der Form $A = QR$ schreiben, wobei $Q \in U_n$ und $R \in M_{n \times n}(\mathbb{C})$ eine (invertierbare) obere Dreiecksmatrix mit positiven (reellen) Diagonaleinträgen ist.

BEWEIS. Wir zeigen nur (b), der erste Teil lässt sich völlig analog beweisen. Da die Matrix $A = (v_1 | \dots | v_n)$ invertierbar ist, bilden ihre Spalten v_1, \dots, v_n eine Basis von \mathbb{C}^n . Wenden wir darauf das Gram-Schmidt Orthonormalisierungsverfahren an, erhalten wir eine Orthonormalbasis b_1, \dots, b_n von \mathbb{C}^n und Skalare $r_{1i}, \dots, r_{ii} \in \mathbb{C}$, sodass

$$v_i = r_{1i} b_1 + \dots + r_{ii} b_i \quad \text{und} \quad r_{ii} > 0, \quad (\text{VII.14})$$

für alle $i = 1, \dots, n$, siehe (VII.12). Die Matrix $Q := (b_1 | \dots | b_n)$ ist unitär, denn $(Q^* Q)_{ij} = b_i^* b_j = \langle b_i, b_j \rangle = \delta_{ij}$, also $Q^* Q = I_n$. Offensichtlich ist

$$R = \begin{pmatrix} r_{11} & \cdots & r_{1n} \\ & \ddots & \vdots \\ & & r_{nn} \end{pmatrix}$$

eine obere Dreiecksmatrix mit positiven reellen Diagonaleinträgen. Die Gleichungen (VII.14) besagen gerade $A = QR$.

Um die Eindeutigkeit der Darstellung einzusehen, seien nun $Q_1, Q_2 \in U_n$ und R_1, R_2 zwei obere Dreiecksmatrizen mit positiven reellen Diagonaleinträgen mit

$$Q_1 R_1 = Q_2 R_2.$$

Es gilt daher

$$R = Q,$$

wobei $Q := (Q_2)^{-1}Q_1 \in U_n$ und $R := R_2(R_1)^{-1}$ eine obere Dreiecksmatrix mit positiven reellen Diagonaleinträgen ist.¹² Daraus erhalten wir $R^*R = Q^*Q = I_n$, also $R^* = R^{-1}$. Beachte, dass R^* eine untere Dreiecksmatrix ist und R^{-1} eine obere Dreiecksmatrix ist. Aus der vorangehenden Gleichheit schließen wir, dass R eine Diagonalmatrix sein muss. Aus $R^*R = I_n$ folgt, dass alle Diagonaleinträge Absolutbetrag 1 haben. Da sie alle positiv und reell sind muss $R = I_n$ gelten, d.h. $R_1 = R_2$. Somit auch $Q = I_n$ und $Q_1 = Q_2$. \square

VII.2.29. BEISPIEL. Wir wollen die QR-Zerlegung der Matrix

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

bestimmen. Wir wenden also das Gram–Schmidt Orthonormalisierungsverfahren auf die Basis

$$v_1 = \begin{pmatrix} 0 \\ 0 \\ 2 \\ 0 \end{pmatrix}, \quad v_2 = \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \quad v_3 = \begin{pmatrix} 1 \\ 0 \\ 3 \\ 3 \end{pmatrix}, \quad v_4 = \begin{pmatrix} 2 \\ 1 \\ -1 \\ 1 \end{pmatrix}$$

an und erhalten:

$$b_1 = \frac{v_1}{\|v_1\|} = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \quad b_2 = \frac{v_2 - \langle b_1, v_2 \rangle b_1}{\|v_2 - \langle b_1, v_2 \rangle b_1\|} = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

$$b_3 = \frac{v_3 - \langle b_1, v_3 \rangle b_1 - \langle b_2, v_3 \rangle b_2}{\|v_3 - \langle b_1, v_3 \rangle b_1 - \langle b_2, v_3 \rangle b_2\|} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}$$

$$b_4 = \frac{v_4 - \langle b_1, v_4 \rangle b_1 - \langle b_2, v_4 \rangle b_2 - \langle b_3, v_4 \rangle b_3}{\|v_4 - \langle b_1, v_4 \rangle b_1 - \langle b_2, v_4 \rangle b_2 - \langle b_3, v_4 \rangle b_3\|} = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}$$

¹²Die Menge der reellen bzw. komplexen oberen Dreiecksmatrizen mit positiven Diagonaleinträgen bildet eine Untergruppe von $GL_n(\mathbb{R})$ bzw. $GL_n(\mathbb{C})$, vgl. Aufgabe 112.

Somit ist $A = QR$ wobei

$$Q = (b_1|b_2|b_3|b_4) = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}, \quad R = Q^{-1}A = Q^t A = \begin{pmatrix} 2 & 1 & 3 & -1 \\ 0 & 1 & 1 & 2 \\ 0 & 0 & 3 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

VII.2.30. DEFINITION (Orthogonales Komplement). Sei V ein Euklidischer oder unitärer Vektorraum und $X \subseteq V$ eine Teilmenge. Unter dem *orthogonalen Komplement* von X verstehen wir den Teilraum

$$X^\perp := \{v \in V \mid \forall x \in X : \langle x, v \rangle = 0\} = \bigcap_{x \in X} \ker(\langle x, - \rangle).$$

VII.2.31. LEMMA. Sei V ein Euklidischer oder unitärer Vektorraum. Weiters seien X, Y, X_1, X_2 Teilmengen und W, W_1, W_2 Teilräume von V . Dann gilt:

- (a) $\{0\}^\perp = V$.
- (b) $X \subseteq X^{\perp\perp}$.
- (c) $X \subseteq Y$ dann $Y^\perp \subseteq X^\perp$.
- (d) $(X_1 \cup X_2)^\perp = X_1^\perp \cap X_2^\perp$.
- (e) $\langle X \rangle^\perp = X^\perp$.
- (f) $W \cap W^\perp = \{0\}$.
- (g) $(W_1 + W_2)^\perp = W_1^\perp \cap W_2^\perp$.

BEWEIS. Die erste Aussage (a) ist trivial. Behauptung (b) folgt sofort aus: $\langle x, v \rangle = 0 \Leftrightarrow \langle v, x \rangle = 0$. Die Aussagen (c) und (d) sind trivial. Ad (e): Die eine Inklusion, $\langle X \rangle^\perp \subseteq X^\perp$, folgt aus (c), denn $X \subseteq \langle X \rangle$. Um auch die umgekehrte Inklusion, $X^\perp \subseteq \langle X \rangle^\perp$, einzusehen, sei $v \in X^\perp$ und $w \in \langle X \rangle$. Es existieren daher $x_i \in X$ und Skalare λ_i , sodass $w = \lambda_1 x_1 + \dots + \lambda_n x_n$. Mit der Bilinearität des inneren Produkts erhalten wir $\langle v, w \rangle = \lambda_1 \langle v, x_1 \rangle + \dots + \lambda_n \langle v, x_n \rangle = 0$, da ja $\langle v, x_i \rangle = 0$. Dies zeigt $v \in \langle X \rangle^\perp$, für jedes $v \in X^\perp$, also $X^\perp \subseteq \langle X \rangle^\perp$. Ad (f): Für $v \in W \cap W^\perp$ gilt $\langle v, v \rangle = 0$ und daher $v = 0$. Behauptung (g) ist eine Konsequenz von (d) und (e), denn $W_1 + W_2 = \langle W_1 \cup W_2 \rangle$ und daher $(W_1 + W_2)^\perp = \langle W_1 \cup W_2 \rangle^\perp = (W_1 \cup W_2)^\perp = W_1^\perp \cap W_2^\perp$. \square

VII.2.32. SATZ (Orthogonalprojektion). Ist V ein Euklidischer oder unitärer Vektorraum und W ein endlich dimensionaler Teilraum von V , dann gilt

$$V = W \oplus W^\perp. \quad (\text{VII.15})$$

Der mit dieser Zerlegung assoziierte Projektor, $p: V \rightarrow W$, wird Orthogonalprojektion auf W genannt. Ist b_1, \dots, b_k eine Orthonormalbasis von W , so gilt

$$p(v) = \sum_{i=1}^k \langle b_i, v \rangle b_i, \quad v \in V. \quad (\text{VII.16})$$

Für jedes $v \in V$ ist $p(v)$ der eindeutig bestimmte Punkt in W mit kleinstem Abstand zu v , d.h. es gilt

$$d(p(v), v) \leq d(w, v),$$

für jedes $w \in W$, und Gleichheit tritt nur dann ein, wenn $w = p(v)$. Der Punkt $p(v) \in W$ mit kleinstem Abstand zu v ist daher durch $v - p(v) \in W^\perp$ eindeutig charakterisiert.

BEWEIS. Sei b_1, \dots, b_k eine Orthonormalbasis von W , vgl. Korollar VII.2.24. Für die durch (VII.16) definierte lineare Abbildung $p: V \rightarrow W$ gilt $p|_W = \text{id}_W$, siehe Proposition VII.2.22, also ist p ein Projektor mit $\text{img}(p) = W$. Für seinen Kern erhalten wir aus der linearen Unabhängigkeit der Vektoren b_1, \dots, b_k ,

$$\ker(p) = \bigcap_{i=1}^k b_i^\perp = \{b_1, \dots, b_k\}^\perp = \langle b_1, \dots, b_k \rangle^\perp = W^\perp,$$

vgl. Lemma VII.2.31(e). Dies zeigt (VII.15) und (VII.16), vgl. Proposition II.5.8. Nach dem Satz von Pythagoras, siehe Proposition VII.2.8, gilt für jedes $w \in W$,

$$\|v - w\|^2 = \|v - p(v) + p(v) - w\|^2 = \|v - p(v)\|^2 + \|p(v) - w\|^2 \geq \|v - p(v)\|^2,$$

denn $v - p(v) \in \ker(p) = W^\perp$ und $p(v) - w \in W$, also $\langle v - p(v), p(v) - w \rangle = 0$. Dies zeigt $d(w, v) \geq d(p(v), v)$. Gleichheit kann nur dann eintreten, wenn $\|p(v) - w\|^2 = 0$, d.h. $w = p(v)$ gilt. \square

VII.2.33. KOROLLAR. Ist V ein endlich dimensionaler Euklidischer oder unitärer Vektorraum dann gilt für jeden Teilraum W von V :

(a) $\dim(W^\perp) = \dim(V) - \dim(W)$.

(b) $W^{\perp\perp} = W$.

(c) $W^\perp = \{0\} \Rightarrow W = V$.

(d) $(W_1 \cap W_2)^\perp = W_1^\perp + W_2^\perp$, für je zwei Teilräume W_1 und W_2 von V .

BEWEIS. Behauptung (a) folgt aus (VII.15) und einer bekannten Dimensionsformel. Daraus erhalten wir auch $\dim(W^{\perp\perp}) = \dim(W)$, also (b), da ja $W \subseteq W^{\perp\perp}$ nach Lemma VII.2.31(b). Behauptung (c) folgt dann aus $\{0\}^\perp = V$, siehe Lemma VII.2.31(a). Aus Lemma VII.2.31(g) erhalten wir

$$(W_1^\perp + W_2^\perp)^\perp = W_1^{\perp\perp} \cap W_2^{\perp\perp} = W_1 \cap W_2$$

und daher $(W_1 \cap W_2)^\perp = (W_1^\perp + W_2^\perp)^{\perp\perp} = W_1^\perp + W_2^\perp$, womit auch (d) gezeigt wäre. \square

Ist W ein endlich dimensionaler Teilraum eines Euklidischen oder unitären Vektorraums V , und ist $v \in V$, dann wird

$$d(v, W) := \min_{w \in W} d(v, w) = d(v, p(v)) = \|v - p(v)\| = \left\| v - \sum_{i=1}^k \langle b_i, v \rangle b_i \right\|$$

der Abstand von v zum Teilraum W genannt. Dabei bezeichnet $p: V \rightarrow W$ die Orthogonalprojektion aus Satz VII.2.32 und b_1, \dots, b_k ist eine beliebige Orthonormalbasis von W . Offensichtlich gilt $d(v, W) = 0$ genau dann, wenn $v \in W$.

VII.2.34. BEISPIEL. Wir wollen die Orthogonalprojektion auf den Teilraum

$$W = \left\langle \begin{pmatrix} 3 \\ 3 \\ 6 \end{pmatrix}, \begin{pmatrix} 3 \\ 7 \\ -2 \end{pmatrix} \right\rangle$$

von \mathbb{R}^3 bestimmen, wobei \mathbb{R}^3 mit dem standard inneren Produkt versehen sei. Wir bestimmen zunächst eine Orthonormalbasis von W in dem wir das Gram-Schmidt Orthonormalisierungsverfahren auf die beiden Vektoren

$$v_1 = \begin{pmatrix} 3 \\ 3 \\ 6 \end{pmatrix} \quad \text{und} \quad v_2 = \begin{pmatrix} 3 \\ 7 \\ -2 \end{pmatrix}$$

anwenden. Dies liefert

$$b_1 = \frac{v_1}{\|v_1\|} = \frac{1}{\sqrt{54}} \begin{pmatrix} 3 \\ 3 \\ 6 \end{pmatrix} = \frac{1}{\sqrt{6}} \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix},$$

$$\tilde{b}_2 = v_2 - \langle b_1, v_2 \rangle b_1 = \begin{pmatrix} 3 \\ 7 \\ -2 \end{pmatrix} - \frac{6}{6} \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix} = \begin{pmatrix} 2 \\ 6 \\ -4 \end{pmatrix} = 2 \begin{pmatrix} 1 \\ 3 \\ -2 \end{pmatrix},$$

$$b_2 = \frac{\tilde{b}_2}{\|\tilde{b}_2\|} = \frac{1}{\sqrt{14}} \begin{pmatrix} 1 \\ 3 \\ -2 \end{pmatrix}.$$

Somit ist b_1, b_2 eine Orthonormalbasis von W , die Orthogonalprojektion auf W ist daher

$$p(v) = \langle b_1, v \rangle b_1 + \langle b_2, v \rangle b_2 = BB^t v, \quad \text{wobei} \quad B = (b_1 | b_2) = \begin{pmatrix} 1/\sqrt{6} & 1/\sqrt{14} \\ 1/\sqrt{6} & 3/\sqrt{14} \\ 2/\sqrt{6} & -2\sqrt{14} \end{pmatrix}.$$

Bezüglich der Standardbasis E von \mathbb{R}^3 gilt daher

$$[p]_{EE} = BB^t = \frac{1}{42} \begin{pmatrix} 10 & 16 & 8 \\ 16 & 34 & -4 \\ 8 & -4 & 40 \end{pmatrix}.$$

VII.3. Normale und selbstadjungierte Operatoren.

VII.3.1. LEMMA. *Sei V ein Euklidischer oder unitärer Vektorraum. Dann definiert*

$$b: V \rightarrow V^*, \quad b(v) := \langle v, - \rangle, \quad (\text{VII.17})$$

eine injektive reell lineare Abbildung, die im unitären Fall darüber hinaus komplex anti-linear ist, d.h. $b(\lambda v) = \bar{\lambda}b(v)$. Ist V endlich dimensional, dann ist (VII.17) ein reell linearer Isomorphismus, insbesondere existiert zu jedem $\alpha \in V^$ ein eindeutiger Vektor $a \in V$, sodass $b(a) = \alpha$, d.h. $\langle a, v \rangle = \alpha(v)$, für alle $v \in V$.*

BEWEIS. Da das innere Produkt linear in der zweiten Eintragung ist, ist (VII.17) wohldefiniert. Die reelle Linearität von b folgt daraus, dass das innere Produkt reell linear in der ersten Eintragung ist. Es gilt $\ker(b) = \{0\}$, denn aus $b(v) = 0$ folgt $0 = b(v)(v) = \langle v, v \rangle = \|v\|^2$ also $v = 0$. Dies zeigt, dass b eine injektive Abbildung ist. Im endlich dimensionalen Fall gilt $\dim_{\mathbb{R}}(V) = \dim_{\mathbb{R}}(V^*)$, nach Korollar IV.2.11 ist b in diesem Fall also auch surjektiv. \square

VII.3.2. PROPOSITION (Adjungierte Abbildung). *Sei $\varphi: V \rightarrow W$ eine lineare Abbildung zwischen endlich dimensionalen Euklidischen oder unitären Vektorräumen. Dann existiert eine eindeutige lineare Abbildung $\varphi^*: W \rightarrow V$, sodass*

$$\langle \varphi^*(w), v \rangle = \langle w, \varphi(v) \rangle, \quad (\text{VII.18})$$

für alle $v \in V$ und $w \in W$. Diese Abbildung φ^ wird als die zu φ adjungierte Abbildung bezeichnet, sie macht folgendes Diagramm kommutativ:*

$$\begin{array}{ccc} W & \xrightarrow{\varphi^*} & V \\ b_W \downarrow \cong & & \cong \downarrow b_V \\ W^* & \xrightarrow{\varphi^t} & V^* \end{array} \quad \text{d.h. es gilt} \quad b_V \circ \varphi^* = \varphi^t \circ b_W,$$

wobei b_V und b_W die (komplex antilinearen) Isomorphismen aus Lemma VII.3.1 und $\varphi^t: W^ \rightarrow V^*$ die duale Abbildung aus Proposition III.4.3 bezeichnen. Für jede weitere lineare Abbildung $\psi: W \rightarrow U$ gilt*

$$\varphi^{**} = \varphi, \quad (\psi \circ \varphi)^* = \varphi^* \circ \psi^*, \quad \text{und} \quad \text{id}_V^* = \text{id}_V.$$

Für zwei lineare Abbildungen $\varphi_1, \varphi_2: V \rightarrow W$ und jeden Skalar λ haben wir

$$(\varphi_1 + \varphi_2)^* = \varphi_1^* + \varphi_2^*, \quad \text{und} \quad (\lambda\varphi)^* = \bar{\lambda}\varphi^*.$$

Sind B und C Orthonormalbasen von V bzw. W , dann gilt

$$[\varphi^*]_{BC} = [\varphi]_{CB}^*. \quad (\text{VII.19})$$

Schließlich ist

$$\ker(\varphi^*) = \text{img}(\varphi)^\perp \quad \text{und} \quad \text{img}(\varphi^*) = \ker(\varphi)^\perp. \quad (\text{VII.20})$$

BEWEIS. Für alle $v \in V$ und $w \in W$ gilt

$$\langle \varphi^*(w), v \rangle = b_V(\varphi^*(w))(v)$$

und

$$\langle w, \varphi(v) \rangle = b_W(w)(\varphi(v)) = (\varphi^t(b_W(w)))(v).$$

Die Bedingung (VII.18) ist daher zu

$$b_V \circ \varphi^* = \varphi^t \circ b_W$$

äquivalent. Somit wird durch $\varphi^* := b_V^{-1} \circ \varphi^t \circ b_W$ eine Abbildung $W \rightarrow V$ definiert, die (VII.18) genügt, und diese ist dadurch eindeutig bestimmt. Als Komposition reell linearer Abbildungen ist φ^* reell linear. Im unitären Fall gilt weiters

$$\begin{aligned} \varphi^*(\lambda v) &= b_V^{-1}(\varphi^t(b_W(\lambda v))) = b_V^{-1}(\varphi^t(\bar{\lambda} b_W(v))) \\ &= b_V^{-1}(\bar{\lambda} \varphi^t(b_W(v))) = \lambda b_V^{-1}(\varphi^t(b_W(v))) = \lambda \varphi^*(v), \end{aligned}$$

d.h. $\varphi^*: W \rightarrow V$ ist auch komplex linear. Aus (VII.18) erhalten wir $\overline{\langle \varphi^*(w), v \rangle} = \langle w, \varphi(v) \rangle$, es gilt daher auch

$$\langle v, \varphi^*(w) \rangle = \langle \varphi(v), w \rangle, \quad v \in V, w \in W. \quad (\text{VII.21})$$

Somit $\langle \varphi^{**}(v), w \rangle = \langle \varphi(v), w \rangle$, für alle $v \in V$ und $w \in W$, also $\varphi^{**} = \varphi$. Die Gleichung $\text{id}_V^* = \text{id}_V$ ist trivial. Aus

$$\begin{aligned} \langle (\psi \circ \varphi)^*(w), v \rangle &= \langle w, (\psi \circ \varphi)(v) \rangle = \langle w, \psi(\varphi(v)) \rangle \\ &= \langle \psi^*(w), \varphi(v) \rangle = \langle \varphi^*(\psi^*(w)), v \rangle = \langle (\varphi^* \circ \psi^*)(w), v \rangle, \end{aligned}$$

erhalten wir $(\psi \circ \varphi)^* = \varphi^* \circ \psi^*$. Analog folgt aus

$$\begin{aligned} \langle (\varphi_1 + \varphi_2)^*(w), v \rangle &= \langle w, (\varphi_1 + \varphi_2)(v) \rangle = \langle w, \varphi_1(v) + \varphi_2(v) \rangle \\ &= \langle w, \varphi_1(v) \rangle + \langle w, \varphi_2(v) \rangle = \langle \varphi_1^*(w), v \rangle + \langle \varphi_2^*(w), v \rangle \\ &= \langle \varphi_1^*(w) + \varphi_2^*(w), v \rangle = \langle (\varphi_1^* + \varphi_2^*)(w), v \rangle \end{aligned}$$

die Gleichung $(\varphi_1 + \varphi_2)^* = \varphi_1^* + \varphi_2^*$. Für jeden Skalar λ gilt

$$\begin{aligned} \langle (\lambda \varphi)^*(w), v \rangle &= \langle w, (\lambda \varphi)(v) \rangle = \langle w, \lambda \varphi(v) \rangle \\ &= \lambda \langle w, \varphi(v) \rangle = \lambda \langle \varphi^*(w), v \rangle = \langle \bar{\lambda} \varphi^*(w), v \rangle = \langle (\bar{\lambda} \varphi^*)(w), v \rangle \end{aligned}$$

also $(\lambda \varphi)^* = \bar{\lambda} \varphi^*$. Sind $B = (b_1, \dots, b_n)$ und $C = (c_1, \dots, c_m)$ Orthonormalbasen von V bzw. W , dann gilt,

$$\varphi(b_i) = \sum_{j=1}^m \langle c_j, \varphi(b_i) \rangle c_j, \quad \text{also} \quad ([\varphi]_{CB})_{ji} = \langle c_j, \varphi(b_i) \rangle,$$

siehe Proposition VII.2.22. Analog haben wir

$$\varphi^*(c_j) = \sum_{i=1}^n \langle b_i, \varphi^*(c_j) \rangle b_i, \quad \text{also} \quad ([\varphi^*]_{BC})_{ij} = \langle b_i, \varphi^*(c_j) \rangle.$$

Es gilt daher

$$([\varphi^*]_{BC})_{ij} = \langle b_i, \varphi^*(c_j) \rangle = \overline{\langle \varphi^*(c_j), b_i \rangle} = \overline{\langle c_j, \varphi(b_i) \rangle} = \overline{([\varphi]_{CB})_{ji}} = ([\varphi^*]_{CB})_{ij},$$

also $[\varphi^*]_{BC} = [\varphi^*]_{CB}$. Für jedes $w \in W$ gilt:

$$\begin{aligned} w \in \ker(\varphi^*) &\Leftrightarrow \varphi^*(w) = 0 \\ &\Leftrightarrow \forall v \in V : \langle \varphi^*(w), v \rangle = 0 \\ &\Leftrightarrow \forall v \in V : \langle w, \varphi(v) \rangle = 0 \\ &\Leftrightarrow \forall u \in \text{img}(\varphi) : \langle w, u \rangle = 0 \\ &\Leftrightarrow w \in \text{img}(\varphi)^\perp, \end{aligned}$$

und somit $\ker(\varphi^*) = \text{img}(\varphi)^\perp$. Wenden wir dies auf φ^* an, erhalten wir

$$\ker(\varphi) = \ker(\varphi^{**}) = \text{img}(\varphi^*)^\perp,$$

und daher auch $\ker(\varphi)^\perp = \text{img}(\varphi^*)^{\perp\perp} = \text{img}(\varphi^*)$, nach Korollar VII.2.33(b). \square

VII.3.3. DEFINITION (Selbstadjungierte Abbildungen). Eine lineare Abbildung $\varphi: V \rightarrow V$ auf einem endlich dimensionalen Euklidischen oder unitären Vektorraum V wird *selbstadjungiert* genannt, falls $\varphi^* = \varphi$ gilt, d.h. wenn

$$\langle \varphi(v), w \rangle = \langle v, \varphi(w) \rangle, \quad \text{für alle } v, w \in V.$$

Im Euklidischen Fall werden selbstadjungierte Abbildungen auch als *symmetrische Abbildungen* bezeichnet.

Eine lineare Abbildung $\varphi: V \rightarrow V$ auf einem endlich dimensionalen unitären Vektorraum V ist genau dann selbstadjungiert, wenn ihre Matrixdarstellung, $[\varphi]_{BB}$, bezüglich einer (und dann jeder) Orthonormalbasis B von V selbstadjungiert ist, d.h. wenn $[\varphi]_{BB}^* = [\varphi]_{BB}$ gilt. Dies folgt sofort aus (VII.19). Insbesondere ist eine Matrix $A \in M_{n \times n}(\mathbb{C})$ genau dann selbstadjungiert, d.h. $A^* = A$, wenn die lineare Abbildung $\mathbb{C}^n \rightarrow \mathbb{C}^n$, $x \mapsto Ax$, bezüglich dem standard inneren Produkt auf \mathbb{C}^n selbstadjungiert ist, denn die Standardbasis ist eine Orthonormalbasis. Die Menge der selbstadjungierten Abbildungen $V \rightarrow V$ bildet einen reellen aber keinen komplexen Teilraum von $L(V, V)$. Die Komposition (nicht kommutierender) selbstadjungierter Abbildungen wird i.A. nicht selbstadjungiert sein.

Eine lineare Abbildung $\varphi: V \rightarrow V$ auf einem endlich dimensionalen Euklidischen Vektorraum V ist genau dann symmetrisch, wenn ihre Matrixdarstellung, $[\varphi]_{BB}$, bezüglich einer (und dann jeder) Orthonormalbasis B von V symmetrisch ist, d.h. wenn $[\varphi]_{BB}^t = [\varphi]_{BB}$ gilt, siehe (VII.19). Insbesondere ist eine Matrix $A \in M_{n \times n}(\mathbb{R})$ genau dann symmetrisch, wenn die lineare Abbildung $\mathbb{R}^n \rightarrow \mathbb{R}^n$, $x \mapsto Ax$, bezüglich dem standard inneren Produkt auf \mathbb{R}^n symmetrisch ist. Die Menge der symmetrischen Abbildungen $V \rightarrow V$ bildet einen Teilraum von $L(V, V)$. Die Komposition (nicht kommutierender) symmetrischer Abbildungen wird i.A. nicht symmetrisch sein.

VII.3.4. BEISPIEL. Ist $\varphi: V \rightarrow V$ eine beliebige lineare Abbildung auf einem Euklidischen oder unitären Vektorraum V , so ist $\varphi\varphi^*: V \rightarrow V$ selbstadjungiert, denn aus den Rechenregeln in Proposition VII.3.2 folgt $(\varphi\varphi^*)^* = \varphi^{**}\varphi^* = \varphi\varphi^*$.

VII.3.5. BEISPIEL. Sei $\varphi: V \rightarrow V$ eine beliebige lineare Abbildung auf einem Euklidischen oder unitären Vektorraum V . Dann ist $\frac{1}{2}(\varphi + \varphi^*): V \rightarrow V$ selbstadjungiert, denn mit den Rechenregeln aus Proposition VII.3.2 erhalten wir:

$$\left(\frac{1}{2}(\varphi + \varphi^*)\right)^* = \frac{1}{2}(\varphi + \varphi^*)^* = \frac{1}{2}(\varphi^* + \varphi^{**}) = \frac{1}{2}(\varphi^* + \varphi) = \frac{1}{2}(\varphi + \varphi^*).$$

Offensichtlich ist

$$\varphi = \frac{1}{2}(\varphi + \varphi^*) + \frac{1}{2}(\varphi - \varphi^*),$$

für die lineare Abbildung $\frac{1}{2}(\varphi - \varphi^*): V \rightarrow V$ gilt

$$\left(\frac{1}{2}(\varphi - \varphi^*)\right)^* = \frac{1}{2}(\varphi - \varphi^*)^* = \frac{1}{2}(\varphi^* - \varphi^{**}) = \frac{1}{2}(\varphi^* - \varphi) = -\frac{1}{2}(\varphi - \varphi^*).$$

Lineare Abbildungen $\psi: V \rightarrow V$, für die $\psi^* = -\psi$ gilt, werden anti-selbstadjungiert oder schiefsymmetrisch genannt. Jede lineare Abbildung $\varphi: V \rightarrow V$ lässt sich daher (in eindeutiger Weise) als Summe einer selbstadjungierten und einer anti-selbstadjungierten linearen Abbildung schreiben. Die Abbildung $\varphi \mapsto \frac{1}{2}(\varphi + \varphi^*)$ ist eine reell lineare Projektion auf den Teilraum der selbstadjungierten linearen Abbildungen, und $\varphi \mapsto \frac{1}{2}(\varphi - \varphi^*)$ ist die komplementäre (reell lineare) Projektion auf den Teilraum der anti-selbstadjungierten linearen Abbildungen.

VII.3.6. BEISPIEL. Ist W ein Teilraum eines endlich dimensionalen Euklidischen oder unitären Vektorraums V , dann ist die Orthogonalprojektion, $p: V \rightarrow W \subseteq V$, siehe Satz VII.2.32, selbstadjungiert. Um dies einzusehen, sei b_1, \dots, b_k eine Orthonormalbasis von $W = \text{img}(p)$ und b_{k+1}, \dots, b_n eine Orthonormalbasis von $\ker(p) = W^\perp$. Es ist dann $B = (b_1, \dots, b_n)$ eine Orthonormalbasis von V , für die $[p]_{BB} = \begin{pmatrix} I_k & 0 \\ 0 & 0 \end{pmatrix}$ gilt. Somit ist $[p]_{BB} = [p]_{BB}^*$ und p also selbstadjungiert. Auch die Spiegelung an W längs W^\perp ,

$$\sigma: V \rightarrow V, \quad \sigma = 2p - \text{id}_V,$$

ist selbstadjungiert, denn $\sigma^* = (2p - \text{id}_V)^* = 2p^* - \text{id}_V^* = 2p - \text{id}_V = \sigma$. Alternativ lässt sich dies auch an der Matrixdarstellung von σ ablesen, denn bezüglich der Orthonormalbasis B oben gilt $[\sigma]_{BB} = \begin{pmatrix} I_k & 0 \\ 0 & -I_{n-k} \end{pmatrix}$, also $[\sigma]_{BB}^* = [\sigma]_{BB}$.

VII.3.7. LEMMA. Sei $\varphi: V \rightarrow V$ eine selbstadjungierte lineare Abbildung auf einem endlich dimensionalen unitären Vektorraum, $\varphi^* = \varphi$. Dann sind alle Eigenwerte von φ reell.

BEWEIS. Sei also $0 \neq v \in V$ ein Eigenvektor zum Eigenwert $\lambda \in \mathbb{C}$, d.h. $\varphi(v) = \lambda v$. Dann gilt

$$\begin{aligned} \lambda \|v\|^2 &= \lambda \langle v, v \rangle = \langle v, \lambda v \rangle = \langle v, \varphi(v) \rangle \\ &= \langle \varphi^*(v), v \rangle = \langle \varphi(v), v \rangle = \langle \lambda v, v \rangle = \bar{\lambda} \langle v, v \rangle = \bar{\lambda} \|v\|^2, \end{aligned}$$

also $\lambda = \bar{\lambda}$, da ja $\|v\|^2 \neq 0$. Dies zeigt, dass λ reell ist. \square

VII.3.8. DEFINITION (Normale Abbildungen). Sei V ein endlich dimensionaler Euklidischer oder unitärer Vektorraum. Eine lineare Abbildung $\varphi: V \rightarrow V$ wird *normal* genannt, wenn $\varphi^*\varphi = \varphi\varphi^*$.

Offensichtlich ist jede selbstadjungierte Abbildung normal. Auch jede anti-selbstadjungierte lineare Abbildung ist normal. Die Summe oder das Produkt zweier (nicht kommutierender) normaler Abbildungen wird i.A. nicht normal sein.

VII.3.9. LEMMA. Für eine lineare Abbildung $\varphi: V \rightarrow V$ auf einem endlich dimensionalen Euklidischen oder unitären Vektorraum sind folgende Aussagen äquivalent:

- (a) φ ist normal, d.h. $\varphi^*\varphi = \varphi\varphi^*$.
- (b) $\langle \varphi(v), \varphi(w) \rangle = \langle \varphi^*(v), \varphi^*(w) \rangle$, für alle $v, w \in V$.
- (c) $\|\varphi(v)\| = \|\varphi^*(v)\|$, für alle $v \in V$.
- (d) Es gilt $[\varphi]_{BB}^*[\varphi]_{BB} = [\varphi]_{BB}[\varphi]_{BB}^*$, für eine (und dann jede) Orthonormalbasis B von V .

BEWEIS. Die Äquivalenz (a) \Leftrightarrow (b) folgt aus

$$\langle (\varphi^*\varphi)(v), w \rangle = \langle \varphi^*(\varphi(v)), w \rangle = \langle \varphi(v), \varphi(w) \rangle$$

und

$$\langle (\varphi\varphi^*)(v), w \rangle = \langle \varphi(\varphi^*(v)), w \rangle = \langle \varphi^*(v), \varphi^*(w) \rangle,$$

vgl. (VII.18) und (VII.21). Die Äquivalenz (b) \Leftrightarrow (c) folgt aus der Polarisierungsidentität, siehe Proposition VII.1.53 bzw. VII.1.22, denn $\langle \varphi(v), \varphi(w) \rangle$ ist eine Hermitesche (symmetrische) Form auf V mit assoziierter quadratischer Form $\|\varphi(v)\|^2$, und $\langle \varphi^*(v), \varphi^*(w) \rangle$ ist eine Hermitesche (symmetrische) Form auf V mit assoziierter quadratischer Form $\|\varphi^*(v)\|^2$. Die Äquivalenz (a) \Leftrightarrow (d) folgt aus

$$[\varphi^*\varphi]_{BB} = [\varphi^*]_{BB}[\varphi]_{BB} = [\varphi]_{BB}^*[\varphi]_{BB}$$

und

$$[\varphi\varphi^*]_{BB} = [\varphi]_{BB}[\varphi^*]_{BB} = [\varphi]_{BB}[\varphi^*]_{BB}^*,$$

vgl. Proposition VII.3.2. □

VII.3.10. LEMMA. Sei V ein endlich dimensionaler Euklidischer oder unitärer Vektorraum und $\varphi: V \rightarrow V$ normal, d.h. $\varphi^*\varphi = \varphi\varphi^*$. Dann gilt:

- (a) $\ker(\varphi) = \ker(\varphi^*)$.
- (b) Ist $0 \neq v \in V$ Eigenvektor zum Eigenwert λ von φ , dann ist v auch Eigenvektor von φ^* mit Eigenwert $\bar{\lambda}$. Es gilt daher $E_\lambda^\varphi = E_{\bar{\lambda}}^{\varphi^*}$.
- (c) Sind $\lambda \neq \mu$ zwei verschiedene Eigenwerte von φ , dann stehen die entsprechenden Eigenräume orthogonal aufeinander, $E_\lambda \perp E_\mu$, d.h. für alle $v \in E_\lambda$ und $w \in E_\mu$ gilt $\langle v, w \rangle = 0$.

BEWEIS. Aus Lemma VII.3.9(c) erhalten wir sofort (a). Um (b) einzusehen, beobachten wir zunächst, dass auch $\varphi - \lambda \text{id}_V: V \rightarrow V$ eine normale Abbildung

ist, denn λid_V ist normal und kommutiert mit φ , mit den Rechenregeln aus Proposition VII.3.2 folgt daher:

$$\begin{aligned} (\varphi - \lambda \text{id}_V)^*(\varphi - \lambda \text{id}_V) &= \varphi^* \varphi - (\lambda \text{id}_V)^* \varphi - \varphi^* (\lambda \text{id}_V) + (\lambda \text{id}_V)^* (\lambda \text{id}_V) \\ &= \varphi \varphi^* - \varphi (\lambda \text{id}_V)^* - (\lambda \text{id}_V) \varphi^* + (\lambda \text{id}_V) (\lambda \text{id}_V)^* \\ &= (\varphi - \lambda \text{id}_V) (\varphi - \lambda \text{id}_V)^*. \end{aligned}$$

Aus (a) erhalten wir nun

$$E_\lambda^\varphi = \ker(\varphi - \lambda \text{id}_V) = \ker((\varphi - \lambda \text{id}_V)^*) = \ker(\varphi^* - \bar{\lambda} \text{id}_V) = E_{\bar{\lambda}}^{\varphi^*},$$

und daher (b). Um (c) einzusehen, seien nun $\lambda \neq \mu$ zwei verschiedene Eigenwerte von φ . Weiters seien $v \in E_\lambda$ und $w \in E_\mu$, d.h. $\varphi(v) = \lambda v$ und $\varphi(w) = \mu w$. Nach (b) gilt daher auch $\varphi^*(v) = \bar{\lambda} v$. Wir erhalten somit

$$\mu \langle v, w \rangle = \langle v, \mu w \rangle = \langle v, \varphi(w) \rangle = \langle \varphi^*(v), w \rangle = \langle \bar{\lambda} v, w \rangle = \bar{\lambda} \langle v, w \rangle = \lambda \langle v, w \rangle,$$

folglich $(\mu - \lambda) \langle v, w \rangle = 0$ und daher $\langle v, w \rangle = 0$, da ja $\mu - \lambda \neq 0$. \square

VII.3.11. SATZ (Spektralsatz für normale Operatoren). *Sei V ein endlich dimensionaler unitärer Vektorraum. Eine lineare Abbildung $\varphi: V \rightarrow V$ ist genau dann normal, wenn eine Orthonormalbasis B von V existiert, sodass $[\varphi]_{BB}$ Diagonalgestalt hat.*

BEWEIS. Sei zunächst B eine Orthonormalbasis von V , sodass $[\varphi]_{BB}$ Diagonalgestalt hat. Dann gilt offensichtlich $[\varphi]_{BB}^* [\varphi]_{BB} = [\varphi]_{BB} [\varphi]_{BB}^*$, nach Lemma VII.3.9 ist φ daher normal.

Sei nun umgekehrt $\varphi: V \rightarrow V$ eine normale lineare Abbildung, d.h. $\varphi^* \varphi = \varphi \varphi^*$. Wir werden nun mittels Induktion nach $\dim(V)$ zeigen, dass eine Orthonormalbasis B von V existiert bezüglich der $[\varphi]_{BB}$ Diagonalgestalt hat. Der Induktionsanfang, $\dim(V) = 0$ ist trivial. Für den Induktionsschritt sei nun $\dim(V) \geq 1$. Da \mathbb{C} algebraisch abgeschlossen ist, existiert ein Eigenwert $\lambda \in \mathbb{C}$ von φ . Offensichtlich ist der Eigenraum E_λ invariant unter φ , d.h. $\varphi(E_\lambda) \subseteq E_\lambda$. Da φ mit φ^* kommutiert, ist E_λ aber auch unter φ^* invariant, denn für jedes $v \in E_\lambda$ gilt $\varphi(\varphi^*(v)) = \varphi^*(\varphi(v)) = \varphi^*(\lambda v) = \lambda \varphi^*(v)$, also $\varphi^*(v) \in E_\lambda$ und somit $\varphi^*(E_\lambda) \subseteq E_\lambda$. Daraus folgt nun, dass auch das orthogonale Komplement, E_λ^\perp , invariant unter φ und φ^* ist: für alle $w \in E_\lambda^\perp$ und $v \in E_\lambda$ gilt nämlich $\langle \varphi^*(w), v \rangle = \langle w, \varphi(v) \rangle = 0$ da $\varphi(v) \in E_\lambda$, also $\varphi^*(w) \in E_\lambda^\perp$ und daher $\varphi^*(E_\lambda^\perp) \subseteq E_\lambda^\perp$; analog haben wir $\langle v, \varphi(w) \rangle = \langle \varphi^*(v), w \rangle = 0$ da $\varphi^*(v) \in E_\lambda$, also $\varphi(w) \in E_\lambda^\perp$ und daher auch $\varphi(E_\lambda^\perp) \subseteq E_\lambda^\perp$. Somit ist

$$V = E_\lambda \oplus E_\lambda^\perp$$

eine unter φ und φ^* invariante orthogonale Zerlegung, vgl. Satz VII.2.32. Für die Adjungierte der Einschränkung, $\varphi|_{E_\lambda^\perp}: E_\lambda^\perp \rightarrow E_\lambda^\perp$, erhalten wir daraus $\varphi|_{E_\lambda^\perp}^* = \varphi^*|_{E_\lambda^\perp}$. Insbesondere ist auch $\varphi|_{E_\lambda^\perp}: E_\lambda^\perp \rightarrow E_\lambda^\perp$ eine normale Abbildung,

$$\varphi|_{E_\lambda^\perp}^* \varphi|_{E_\lambda^\perp} = \varphi^*|_{E_\lambda^\perp} \varphi|_{E_\lambda^\perp} = (\varphi^* \varphi)|_{E_\lambda^\perp} = (\varphi \varphi^*)|_{E_\lambda^\perp} = \varphi|_{E_\lambda^\perp} \varphi^*|_{E_\lambda^\perp} = \varphi|_{E_\lambda^\perp} \varphi|_{E_\lambda^\perp}^*.$$

Nach Induktionsvoraussetzung existiert daher eine Orthonormalbasis B'' von E_λ^\perp , sodass $[\varphi|_{E_\lambda^\perp}]_{B''B''}$ Diagonalgestalt hat, denn $\dim(E_\lambda^\perp) = \dim(V) - \dim(E_\lambda) < \dim(V)$, siehe Korollar VII.2.33(a). Bezeichnet B' eine beliebige Orthonormalbasis von E_λ , so gilt $[\varphi|_{E_\lambda}]_{B'B'} = \lambda I_k$, wobei $k = \dim(E_\lambda)$, denn $\varphi|_{E_\lambda} = \lambda \operatorname{id}_{E_\lambda}$. Folglich ist $B = B' \cup B''$ eine Orthonormalbasis von V , und

$$[\varphi]_{BB} = \begin{pmatrix} [\varphi|_{E_\lambda}]_{B'B'} & 0 \\ 0 & [\varphi|_{E_\lambda^\perp}]_{B''B''} \end{pmatrix} = \begin{pmatrix} I_k & 0 \\ 0 & [\varphi|_{E_\lambda^\perp}]_{B''B''} \end{pmatrix}$$

hat Diagonalgestalt. \square

VII.3.12. KOROLLAR (Spektralzerlegung). *Sei $\varphi: V \rightarrow V$ eine normale Abbildung auf einem endlich dimensionalen unitären Vektorraum V . Dann ist φ diagonalisierbar und Eigenräume zu verschiedenen Eigenwerten stehen orthogonal aufeinander. Bezeichnen $\lambda_1, \dots, \lambda_k$ die Eigenwerte von φ , dann zerfällt V in eine orthogonale direkte Summe der Eigenräume,*

$$V = E_{\lambda_1} \oplus \dots \oplus E_{\lambda_k},$$

und es gilt

$$\varphi = \lambda_1 \pi_1 + \dots + \lambda_k \pi_k,$$

wobei $\pi_i: V \rightarrow E_{\lambda_i} \subseteq V$ die Orthogonalprojektion auf den Eigenraum zum Eigenwert λ_i bezeichnet. Diese Projektionen genügen den Relationen:

$$\operatorname{id}_V = \pi_1 + \dots + \pi_k, \quad \pi_i^* = \pi_i = \pi_i^2, \quad \pi_i \pi_j = 0 = \pi_j \pi_i, \quad i \neq j.$$

Insbesondere gilt dies alles für selbstadjungierte φ , in diesem Fall sind darüber hinaus alle Eigenwerte reell.

BEWEIS. Dies folgt sofort aus Satz VII.3.11, siehe auch Lemma VII.3.10(c), Satz VI.1.18, Proposition VI.1.16(e), Aufgabe 115 und Lemma VII.3.7. \square

Für Matrizen erhalten wir daraus:

VII.3.13. KOROLLAR. *Sei $A \in M_{n \times n}(\mathbb{C})$ eine normale Matrix, d.h. $A^*A = AA^*$. Dann existiert eine unitäre Matrix $U \in U_n$, sodass $U^{-1}AU = U^*AU$ Diagonalgestalt hat. Insbesondere sind normale Matrizen diagonalisierbar und Eigenräume zu verschiedenen Eigenwerten stehen orthogonal aufeinander. Ist A selbstadjungiert, dann sind darüber hinaus alle Eigenwerte von A , und auch die Einträge der Diagonalmatrix $U^{-1}AU$, reell.*

BEWEIS. Betrachte \mathbb{C}^n mit dem standard inneren Produkt, $\langle x, y \rangle = x^*y$. Nach Lemma VII.3.9 ist $\varphi: \mathbb{C}^n \rightarrow \mathbb{C}^n$, $\varphi(x) := Ax$, eine normale Abbildung. Nach Satz VII.3.11 existiert daher eine Orthonormalbasis $B = (b_1, \dots, b_n)$ von \mathbb{C}^n , sodass $[\varphi]_{BB}$ eine Diagonalmatrix ist. Die unitäre Matrix $U := (b_1 | \dots | b_n)$ hat daher die gewünschte Eigenschaft. Bezeichnet nämlich E die Standardbasis so gilt $[\varphi]_{EE} = A$, $U = T_{EB}$ und $U^{-1}AU = T_{EB}^{-1}[\varphi]_{EE}T_{EB} = [\varphi]_{BB}$. \square

VII.3.14. SATZ (Spektralsatz für symmetrische Operatoren). *Sei V ein endlich dimensionaler Euklidischer Vektorraum. Eine lineare Abbildung $\varphi: V \rightarrow V$ ist genau dann selbstadjungiert (symmetrisch) wenn eine Orthonormalbasis B von V existiert, sodass $[\varphi]_{BB}$ Diagonalgestalt hat.*

BEWEIS. Wir zeigen zunächst, dass jede symmetrische Abbildung $\varphi: V \rightarrow V$ einen (reellen) Eigenwert besitzt. Sei dazu B eine Orthonormalbasis von V . Dann ist $A = [\varphi]_{BB} \in M_{n \times n}(\mathbb{R})$ eine symmetrische Matrix, $A^t = A$. Fassen wir A als komplexe Matrix auf, $A \in M_{n \times n}(\mathbb{C})$, so gilt $A^* = A$. Nach Korollar VII.3.13 besitzt A daher einen reellen Eigenwert, also hat auch $\varphi: V \rightarrow V$ einen reellen Eigenwert. Damit lässt sich der Satz nun völlig analog zum Beweis von Satz VII.3.11 zeigen. \square

VII.3.15. KOROLLAR (Spektralzerlegung). *Sei $\varphi: V \rightarrow V$ eine selbstadjungierte (symmetrische) Abbildung auf einem endlich dimensionalen Euklidischen Vektorraum V . Dann ist φ diagonalisierbar und Eigenräume zu verschiedenen Eigenwerten stehen orthogonal aufeinander. Bezeichnen $\lambda_1, \dots, \lambda_k$ die (reellen) Eigenwerte von φ , dann zerfällt V in eine orthogonale direkte Summe der Eigenräume,*

$$V = E_{\lambda_1} \oplus \dots \oplus E_{\lambda_k},$$

und es gilt

$$\varphi = \lambda_1 \pi_1 + \dots + \lambda_k \pi_k,$$

wobei $\pi_i: V \rightarrow E_{\lambda_i} \subseteq V$ die Orthogonalprojektion auf den Eigenraum zum Eigenwert λ_i bezeichnet. Diese Projektionen genügen den Relationen:

$$\text{id}_V = \pi_1 + \dots + \pi_k, \quad \pi_i^* = \pi_i = \pi_i^2, \quad \pi_i \pi_j = 0 = \pi_j \pi_i, \quad i \neq j.$$

VII.3.16. KOROLLAR. *Sei $A \in M_{n \times n}(\mathbb{R})$ symmetrisch, d.h. $A^t = A$. Dann existiert eine orthogonale Matrix $U \in O_n$, sodass $U^{-1}AU = U^tAU$ Diagonalgestalt hat. Insbesondere sind reelle symmetrische Matrizen stets diagonalisierbar und Eigenräume zu verschiedenen Eigenwerten stehen orthogonal aufeinander.*

VII.3.17. BEISPIEL. Die symmetrische reelle Matrix

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

hat charakteristisches Polynom $p = \det(A - zI_3) = -z^3 + 3z + 2 = -(z+1)^2(z-2)$. Nach Korollar VII.3.16 existiert daher eine orthogonale Matrix $U \in O_3$, sodass

$$U^{-1}AU = U^tAU = \begin{pmatrix} -1 & & \\ & -1 & \\ & & 2 \end{pmatrix}. \quad (\text{VII.22})$$

Wir wollen nun eine solche Matrix U bestimmen. Für die Eigenräume erhalten wir zunächst

$$E_{-1} = \ker(A + I_3) = \ker \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix} = \left\langle \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix} \right\rangle$$

und

$$E_2 = \ker(A - 2I_3) = \ker \begin{pmatrix} -2 & 1 & 1 \\ 1 & -2 & 1 \\ 1 & 1 & -2 \end{pmatrix} = \left\langle \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \right\rangle.$$

Durch Anwenden des Gram-Schmidt Orthonormalisierungsverfahrens erhalten wir Orthonormalbasen der Eigenräume,

$$E_{-1} = \left\langle \begin{pmatrix} 1/\sqrt{2} \\ -1/\sqrt{2} \\ 0 \end{pmatrix}, \begin{pmatrix} 1/\sqrt{6} \\ 1/\sqrt{6} \\ -2/\sqrt{6} \end{pmatrix} \right\rangle \quad \text{und} \quad E_2 = \left\langle \begin{pmatrix} 1/\sqrt{3} \\ 1/\sqrt{3} \\ 1/\sqrt{3} \end{pmatrix} \right\rangle.$$

Folglich ist

$$U = \begin{pmatrix} 1/\sqrt{2} & 1/\sqrt{6} & 1/\sqrt{3} \\ -1/\sqrt{2} & 1/\sqrt{6} & 1/\sqrt{3} \\ 0 & -2/\sqrt{6} & 1/\sqrt{3} \end{pmatrix}$$

eine orthogonale Matrix, die (VII.22) erfüllt.

VII.3.18. BEISPIEL. Normale reelle Matrizen sind i.A. nicht diagonalisierbar. Etwa ist die reelle Matrix

$$\begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix}$$

normal aber nicht selbstadjungiert, $\theta \in \mathbb{R} \setminus \pi\mathbb{Z}$. Da ihre Eigenwerte, $\cos \theta \pm i \sin \theta$, nicht reell sind kann sie über \mathbb{R} nicht diagonalisierbar sein.

VII.4. Isometrien. Isometrien sind Längen und Winkel bewahrende lineare Abbildungen.

VII.4.1. DEFINITION (Isometrien). Eine lineare Abbildung zwischen Euklidischen oder unitären Vektorräumen, $\varphi: V \rightarrow W$, wird *Isometrie* genannt, falls

$$\langle \varphi(v_1), \varphi(v_2) \rangle = \langle v_1, v_2 \rangle,$$

für alle $v_1, v_2 \in V$.

Offensichtlich ist die Komposition von Isometrien wieder eine Isometrie. Beachte auch, dass Isometrien stets injektiv sind, denn aus $\varphi(v) = 0$ folgt $0 = \langle \varphi(v), \varphi(v) \rangle = \langle v, v \rangle = \|v\|^2$, also $v = 0$.

VII.4.2. LEMMA. Für eine lineare Abbildung $\varphi: V \rightarrow W$ zwischen endlich dimensional Euklidischen oder unitären Vektorräumen sind folgende Aussagen äquivalent:

(a) φ ist eine Isometrie.

- (b) $\varphi^* \varphi = \text{id}_V$.
 (c) $\|\varphi(v)\| = \|v\|$, für alle $v \in V$.
 (d) Für eine (und dann jede) Orthonormalbasis b_1, \dots, b_n von V bilden die Vektoren $\varphi(b_1), \dots, \varphi(b_n)$ ein Orthonormalsystem in W .

Ist $\dim(V) = \dim(W)$, so sind diese Bedingungen auch zu folgender äquivalent:

- (e) $\varphi \varphi^* = \text{id}_W$.

BEWEIS. Aus $\langle \varphi(v_1), \varphi(v_2) \rangle = \langle (\varphi^* \varphi)(v_1), v_2 \rangle$ erhalten wir sofort die Äquivalenz (a) \Leftrightarrow (b). Aus der Polarisierungsidentität folgt die Äquivalenz (a) \Leftrightarrow (c), denn $\langle \varphi(v_1), \varphi(v_2) \rangle$ ist eine Hermitesche/symmetrische Form auf V mit assoziierter quadratischer Form $\|\varphi(v)\|^2$, und $\langle v_1, v_2 \rangle$ ist eine Hermitesche/symmetrische Form auf V mit assoziierter quadratischer Form $\|v\|^2$. Die Implikation (a) \Rightarrow (d) ist offensichtlich. Um auch (d) \Rightarrow (b) einzusehen, sei nun $B = (b_1, \dots, b_n)$ eine Orthonormalbasis von V , sodass $\tilde{C} = (\varphi(b_1), \dots, \varphi(b_n))$ ein Orthonormalsystem in W bildet. Ergänzen wir \tilde{C} zu einer Orthonormalbasis C von W , dann gilt also $[\varphi]_{CB} = \begin{pmatrix} I_n \\ 0 \end{pmatrix}$. Wir erhalten

$$[\varphi^* \varphi]_{BB} = [\varphi^*]_{BC} [\varphi]_{CB} = [\varphi]_{CB}^* [\varphi]_{CB} = (I_n | 0) \begin{pmatrix} I_n \\ 0 \end{pmatrix} = I_n = [\text{id}_V]_{BB},$$

also $\varphi^* \varphi = \text{id}_V$. Damit ist die Äquivalenz der ersten vier Aussagen gezeigt.

Sei nun $\dim(V) = \dim(W)$ und $\varphi: V \rightarrow W$ eine Isometrie. Da Isometrien stets injektiv sind, folgt aus Dimensionsgründen, dass φ ein Isomorphismus mit Inverser $\varphi^{-1} = \varphi^*$ ist. Es gilt daher auch $\varphi \varphi^* = \varphi \varphi^{-1} = \text{id}_W$. Dies zeigt die Implikation (a) \Rightarrow (e). Analog folgt aus $\varphi \varphi^* = \text{id}_W$ zunächst $\varphi^{-1} = \varphi^*$ und dann $\varphi^* \varphi = \varphi^{-1} \varphi = \text{id}_V$. Damit ist auch (e) \Rightarrow (b) gezeigt. \square

Nach dem vorangehenden Lemma bildet die Menge der Isometrien eines endlich-dimensionalen unitären Vektorraums V ,

$$U(V) := \{\varphi: V \rightarrow V \mid \varphi^* \varphi = \text{id}_V\} = \{\varphi: V \rightarrow V \mid \varphi \varphi^* = \text{id}_V\},$$

bezüglich Komposition von Abbildungen eine Gruppe, die die *unitäre Gruppe* des unitären Vektorraums V genannt wird. Wir werden Isometrien $\varphi: V \rightarrow V$ auch als *unitäre* Abbildungen bezeichnen. Eine lineare Abbildung $\varphi: V \rightarrow V$ ist genau dann unitär, wenn $[\varphi]_{BB}$ eine unitäre Matrix ist, bezüglich einer (und dann jeder) Orthonormalbasis B von V . Jede Orthonormalbasis B von V liefert einen Gruppenisomorphismus

$$U(V) \cong U_n, \quad \varphi \leftrightarrow [\varphi]_{BB},$$

wobei $n = \dim(V)$ und

$$U_n = \{A \in M_{n \times n}(\mathbb{C}) \mid A^* A = I_n\} = \{A \in M_{n \times n}(\mathbb{C}) \mid A A^* = I_n\}.$$

Beachte, dass unitäre Abbildungen stets normal sind.

Analog bildet die Menge der Isometrien eines endlich dimensionalen Euklidischen Vektorraums V ,

$$\mathrm{O}(V) := \{\varphi: V \rightarrow V \mid \varphi^* \varphi = \mathrm{id}_V\} = \{\varphi: V \rightarrow V \mid \varphi \varphi^* = \mathrm{id}_V\},$$

eine Gruppe, die als *orthogonale Gruppe* von V bezeichnet wird. Wir werden Isometrien $\varphi: V \rightarrow V$ auch als *orthogonale Abbildungen* bezeichnen. Eine lineare Abbildung $\varphi: V \rightarrow V$ ist genau dann orthogonal, wenn $[\varphi]_{BB}$ eine orthogonale Matrix ist, bezüglich einer (und dann jeder) Orthonormalbasis B von V . Jede Orthonormalbasis B von V liefert einen Gruppenisomorphismus

$$\mathrm{O}(V) \cong \mathrm{O}_n, \quad \varphi \leftrightarrow [\varphi]_{BB},$$

wobei $n = \dim(V)$ und

$$\mathrm{O}_n = \{A \in M_{n \times n}(\mathbb{R}) \mid A^t A = I_n\} = \{A \in M_{n \times n}(\mathbb{R}) \mid A A^t = I_n\}.$$

Beachte, dass orthogonale Abbildungen stets normal sind.

VII.4.3. BEISPIEL (Spiegelungen). Sei V ein endlich dimensionaler Euklidischer oder unitärer Vektorraum und $W \subseteq V$ ein Teilraum. Es bezeichne $\sigma: V \rightarrow V$, die Spiegelung an W längs W^\perp , d.h. $\sigma = 2p - \mathrm{id}_V$, wobei p die Orthogonalprojektion auf W bezeichnet. Wir haben weiter oben bereits gesehen, dass diese Spiegelung selbstadjungiert ist, $\sigma^* = \sigma$. Offensichtlich gilt aber auch $\sigma^2 = \mathrm{id}_V$, und daher $\sigma^* \sigma = \mathrm{id}_V$. Jede solche orthogonale Spiegelung ist daher eine Isometrie, also unitär bzw. orthogonal.

VII.4.4. BEISPIEL (Drehungen). Sei V ein endlich dimensionaler Euklidischer Vektorraum, $E \subseteq V$ ein 2-dimensionaler orientierter Teilraum und $\theta \in \mathbb{R}$. Weiters sei b_1, b_2 eine positiv orientierte Orthonormalbasis von E . Wir ergänzen zu einer Orthonormalbasis $B = (b_1, b_2, b_3, \dots, b_n)$ von V und definieren eine lineare Abbildung

$$\rho_E^\theta: V \rightarrow V, \quad \text{durch} \quad [\rho_E^\theta]_{BB} := \begin{pmatrix} \cos \theta & -\sin \theta & & \\ \sin \theta & \cos \theta & & \\ & & & \\ & & & I_{n-2} \end{pmatrix}.$$

Es lässt sich zeigen, dass dies nicht von der Wahl der Orthonormalbasis abhängt, vgl. Aufgabe 121. Offensichtlich ist ρ_E^θ eine orthogonale Abbildung, $(\rho_E^\theta)^* \rho_E^\theta = \mathrm{id}_V$, denn die definierende Matrix ist orthogonal. Die Abbildung ρ_E^θ ist eine Drehung in E um den Winkel θ , die Punkte in E^\perp werden festgelassen.

VII.4.5. KOROLLAR. Sei V ein endlich dimensionaler unitärer Vektorraum und $\varphi: V \rightarrow V$ unitär. Dann existiert eine Orthonormalbasis B von V , sodass $[\varphi]_{BB}$ eine Diagonalmatrix ist, deren Diagonaleinträge alle Absolutbetrag Eins haben. Insbesondere ist φ diagonalisierbar, alle Eigenwerte haben Absolutbetrag Eins, und Eigenräume zu verschiedenen Eigenwerten stehen normal aufeinander.

BEWEIS. Sei $\lambda \in \mathbb{C}$ ein Eigenwert von φ und $0 \neq v \in V$ ein entsprechender Eigenvektor, d.h. $\varphi(v) = \lambda v$. Mit Lemma VII.4.2 folgt

$$|\lambda| \|v\| = \|\lambda v\| = \|\varphi(v)\| = \|v\|,$$

also $|\lambda| = 1$. Somit haben alle Eigenwerte einer unitären Abbildung Absolutbetrag Eins. Da unitäre Abbildungen auch normal sind, $\varphi^* \varphi = \text{id}_V = \varphi \varphi^*$, folgt das Korollar daher aus Satz VII.3.11 bzw. Korollar VII.3.12. \square

Orthogonale Abbildungen sind i.A. nicht diagonalisierbar, etwa ist die durch die Matrix

$$A = \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix}$$

definierte Drehung, $\mathbb{R}^2 \rightarrow \mathbb{R}^2$, $x \mapsto Ax$, für $\theta \in \mathbb{R} \setminus \mathbb{Z}$ nicht diagonalisierbar, da ihre Eigenwerte, $\cos \theta \pm \mathbf{i} \sin \theta$, nicht reell sind.

Für jeden Einheitsvektor $a \in V$, $\|a\| = 1$, eines Euklidischen Vektorraums bezeichnen wir mit $\sigma_a: V \rightarrow V$ die Spiegelung an der Hyperebene a^\perp längs $\langle a \rangle$,

$$\sigma_a: V \rightarrow V, \quad \sigma_a(v) = v - 2\langle a, v \rangle a, \quad v \in V.$$

Dies ist eine orthogonale Abbildung, $\sigma_a \in O(V)$, denn $\sigma_a^2 = \text{id}_V$ und $\sigma_a^* = \sigma_a$.

VII.4.6. SATZ. Sei $\varphi: V \rightarrow V$ eine orthogonale Abbildung auf einem n -dimensionalen Euklidischen Vektorraum, d.h. $\varphi^* \varphi = \text{id}_V$. Dann existieren $0 \leq k \leq n$ und normierte Vektoren $a_1, \dots, a_k \in V$, sodass

$$\varphi = \sigma_{a_1} \circ \dots \circ \sigma_{a_k}$$

d.h. φ lässt sich als Komposition von höchstens n orthogonalen Spiegelungen an Hyperebenen schreiben.

BEWEIS. Wir führen den Beweis mittels Induktion nach n . Der Induktionsanfang, $n = 0$, ist trivial. Für den Induktionsschritt dürfen wir o.B.d.A. $\varphi \neq \text{id}_V$ annehmen. Es existiert daher $0 \neq w \in V$, sodass $\varphi(w) \neq w$. Die mit dem Einheitsvektor

$$a := \frac{\varphi(w) - w}{\|\varphi(w) - w\|}$$

assoziierte Spiegelung genügt $\sigma_a(w) = \varphi(w)$, denn aus

$$\langle w + \varphi(w), w - \varphi(w) \rangle = \underbrace{\|w\|^2 - \|\varphi(w)\|^2}_{=0} - \underbrace{\langle w, \varphi(w) \rangle + \langle \varphi(w), w \rangle}_{=0} = 0$$

erhalten wir

$$w = \frac{1}{2}(w + \varphi(w)) + \frac{1}{2}(w - \varphi(w)),$$

wobei $\frac{1}{2}(w + \varphi(w)) \in a^\perp$ und $\frac{1}{2}(w - \varphi(w)) \in \langle a \rangle$, also

$$\begin{aligned}\sigma_a(w) &= \sigma_a\left(\frac{1}{2}(w + \varphi(w)) + \frac{1}{2}(w - \varphi(w))\right) \\ &= \sigma_a\left(\frac{1}{2}(w + \varphi(w))\right) + \sigma_a\left(\frac{1}{2}(w - \varphi(w))\right) \\ &= \frac{1}{2}(w + \varphi(w)) - \frac{1}{2}(w - \varphi(w)) \\ &= \varphi(w).\end{aligned}$$

Somit ist $\psi := \sigma_a^{-1}\varphi \in O(V)$ und $\psi(w) = w$. Wegen der Orthogonalität von ψ ist auch die Hyperebene $W := w^\perp$ invariant unter ψ , und die Einschränkung, $\psi|_W: W \rightarrow W$, ist offensichtlich wieder orthogonal, d.h. $\psi|_W \in O(W)$. Nach Induktionsvoraussetzung existieren $0 \leq k \leq n$ und $a_2, \dots, a_k \in W$, sodass $\psi|_W = \sigma_{a_2}^W \cdots \sigma_{a_k}^W$, wobei $\sigma_{a_i}^W: W \rightarrow W$ die Spiegelungen in W bezeichnen. Da $\sigma_{a_i}|_W = \sigma_{a_i}^W$, $\sigma_{a_i}(w) = w$ und $\psi(w) = w$ gilt daher auch $\psi = \sigma_{a_2} \cdots \sigma_{a_k}$ und wir erhalten die gewünschte Darstellung, $\varphi = \sigma_a \psi = \sigma_a \sigma_{a_2} \cdots \sigma_{a_k}$. \square

VII.4.7. SATZ. *Sei V ein endlich dimensionaler Euklidischer Vektorraum und $f: V \rightarrow V$ eine beliebige Abbildung, sodass $d(f(v), f(w)) = d(v, w)$, für alle $v, w \in V$. Dann ist f eine affine Isometrie, d.h. es existiert $\varphi \in O(V)$ und $b \in V$, sodass $f(v) = \varphi(v) + b$, für alle $v \in V$.*

BEWEIS. Sei $b := f(0)$ und $\varphi: V \rightarrow V$, $\varphi(v) := f(v) - b$. Es genügt zu zeigen, dass φ eine lineare orthogonale Abbildung ist. Nach Voraussetzung an f gilt

$$\|\varphi(w) - \varphi(v)\| = \|w - v\|, \quad \text{für alle } v, w \in V, \quad (\text{VII.23})$$

denn $\|\varphi(w) - \varphi(v)\| = \|f(w) - f(v)\| = d(f(v), f(w)) = d(v, w) = \|w - v\|$. Da $\varphi(0) = 0$, gilt daher auch

$$\|\varphi(v)\| = \|v\|, \quad \text{für alle } v \in V. \quad (\text{VII.24})$$

Die Polarisierungsidentität in Proposition VII.2.3 lässt sich in der Form

$$-2\langle v, w \rangle = \|w - v\|^2 - \|v\|^2 - \|w\|^2,$$

schreiben, aus (VII.23) und (VII.24) erhalten wir daher

$$\langle \varphi(v), \varphi(w) \rangle = \langle v, w \rangle, \quad \text{für alle } v, w \in V, \quad (\text{VII.25})$$

denn $-2\langle \varphi(v), \varphi(w) \rangle = \|\varphi(w) - \varphi(v)\|^2 - \|\varphi(v)\|^2 - \|\varphi(w)\|^2 = \|w - v\|^2 - \|v\|^2 - \|w\|^2 = -2\langle v, w \rangle$. Sei nun b_1, \dots, b_n eine Orthonormalbasis von V . Nach (VII.25) ist daher auch $\varphi(b_1), \dots, \varphi(b_n)$ eine Orthonormalbasis von V . Es gilt daher

$$\varphi(v) = \sum_{i=1}^n \langle \varphi(b_i), \varphi(v) \rangle \varphi(b_i) = \sum_{i=1}^n \langle b_i, v \rangle \varphi(b_i),$$

folglich ist φ eine lineare Abbildung. Da φ die Orthonormalbasis b_1, \dots, b_n auf die Orthonormalbasis $\varphi(b_1), \dots, \varphi(b_n)$ abbildet, ist φ auch orthogonal. \square

VII.5. Positivität.

VII.5.1. DEFINITION (Positivität). Eine *selbstadjungierte* lineare Abbildung $\varphi: V \rightarrow V$ auf einem Euklidischen oder unitären Vektorraum V heißt *semipositiv*, falls

$$\langle v, \varphi(v) \rangle \geq 0,$$

für alle $v \in V$. Wir schreiben in diesem Fall $\varphi \geq 0$. Gilt sogar

$$\langle v, \varphi(v) \rangle > 0,$$

für alle $0 \neq v \in V$, so wird φ *positiv* genannt und wir schreiben $\varphi > 0$.

VII.5.2. LEMMA. Für eine lineare Abbildung $\varphi: V \rightarrow V$ auf einem endlich dimensionalen Euklidischen Vektorraum V sind äquivalent:

- (a) φ ist (semi)positiv.
- (b) $\langle v, \varphi(w) \rangle$ ist eine positiv (semi)definite symmetrische Bilinearform auf V .
- (c) Die Matrix $[\varphi]_{BB}$ ist positiv (semi)definit, für eine (und dann jede) Orthonormalbasis B von V .

Für eine lineare Abbildung $\varphi: V \rightarrow V$ auf einem endlich dimensionalen unitären Vektorraum V sind äquivalent:

- (a) φ ist (semi)positiv.
- (b) $\langle v, \varphi(w) \rangle$ ist eine positiv (semi)definite Hermitesche Form auf V .
- (c) Die Matrix $[\varphi]_{BB}$ ist positiv (semi)definit, für eine (und dann jede) Orthonormalbasis B von V .

BEWEIS. Wir werden dies nur für unitäre Vektorräume V zeigen, der Euklidische Fall lässt sich völlig analog behandeln. Eine lineare Abbildung $\varphi: V \rightarrow V$ ist genau dann selbstadjungiert, wenn $h(v, w) = \langle v, \varphi(w) \rangle$ eine Hermitesche Form auf V definiert, denn offensichtlich ist $\langle v, \varphi(w) \rangle$ linear in w und es gilt $\overline{h(w, v)} = \overline{\langle w, \varphi(v) \rangle} = \langle \varphi(v), w \rangle = \langle v, \varphi^*(w) \rangle$. Daraus erhalten wir sofort die Äquivalenz (a) \Leftrightarrow (b). Ist $B = (b_1, \dots, b_n)$ eine Orthonormalbasis von V , dann gilt $[h]_B = [\varphi]_{BB}$, denn nach Proposition VII.2.22 haben wir

$$\varphi(b_j) = \sum_{i=1}^n \langle b_i, \varphi(b_j) \rangle b_i = \sum_{i=1}^n h(b_i, b_j) b_i,$$

und daher $([\varphi]_{BB})_{ij} = h(b_i, b_j) = ([h]_B)_{ij}$. Dies zeigt die Äquivalenz (b) \Leftrightarrow (c), vgl. Definition VII.1.29. \square

VII.5.3. BEISPIEL. Für $a \in M_{1 \times 1}(\mathbb{C}) = \mathbb{C}$ betrachte die lineare Abbildung $\psi_a: \mathbb{C} \rightarrow \mathbb{C}$, $\psi_a(x) = ax$. Diese Abbildung ist genau dann selbstadjungiert, wenn $a \in \mathbb{R}$. Sie ist genau dann (semi)positiv, wenn $a > 0$ bzw. $a \geq 0$ gilt. Die Abbildung ψ_a ist genau dann unitär, wenn $|a| = 1$.

VII.5.4. BEISPIEL. Die lineare Abbildung

$$\psi: \mathbb{R}^3 \rightarrow \mathbb{R}^3, \quad \psi(x) = Ax, \quad \text{wobei} \quad A = \begin{pmatrix} 3 & 1 & 1 \\ 1 & 4 & 1 \\ 1 & 1 & 5 \end{pmatrix}$$

ist positiv (bezüglich des standard inneren Produkts auf \mathbb{R}^3), denn die Matrix A ist symmetrisch und positiv definit, da

$$3 > 0, \quad \begin{vmatrix} 3 & 1 \\ 1 & 4 \end{vmatrix} = 11 > 0, \quad \begin{vmatrix} 3 & 1 & 1 \\ 1 & 4 & 1 \\ 1 & 1 & 5 \end{vmatrix} = 50 > 0,$$

vgl. das Sylvesterkriterium in Satz VII.1.42.

VII.5.5. LEMMA. Seien φ , φ_1 und φ_2 lineare Abbildungen auf einem endlich dimensionalen Euklidischen oder unitären Vektorraum V und $\lambda \in \mathbb{R}$. Dann gilt:

- (a) Ist $\varphi_1 \geq 0$ und $\varphi_2 \geq 0$, dann auch $\varphi_1 + \varphi_2 \geq 0$.
- (b) Ist $\varphi_1 \geq 0$ und $\varphi_2 > 0$, dann auch $\varphi_1 + \varphi_2 > 0$.
- (c) Ist $\lambda \geq 0$ und $\varphi \geq 0$, dann auch $\lambda\varphi \geq 0$.
- (d) Ist $\lambda > 0$ und $\varphi > 0$, dann auch $\lambda\varphi > 0$.
- (e) Ist $\varphi \geq 0$ und $\psi: V \rightarrow W$ linear, dann auch $\psi\varphi\psi^* \geq 0$.
- (f) Ist $\varphi > 0$ und $\psi: V \rightarrow W$ surjektiv und linear, dann auch $\psi\varphi\psi^* > 0$.
- (g) Ist $\varphi \geq 0$ und $-\varphi \geq 0$, dann gilt schon $\varphi = 0$.
- (h) Es gilt $\varphi > 0$ genau dann wenn $\varphi \geq 0$ und φ injektiv (invertierbar) ist.

Insbesondere ist die Menge der (semi)positiven Abbildungen auf V konvex.

BEWEIS. Ist $\varphi_1 \geq 0$ und $\varphi_2 \geq 0$, dann ist $\varphi_1 + \varphi_2$ als Summe zweier selbstadjungierter Abbildungen wieder selbstadjungiert und es gilt $\langle v, (\varphi_1 + \varphi_2)(v) \rangle = \langle v, \varphi_1(v) \rangle + \langle v, \varphi_2(v) \rangle \geq 0$, also $\varphi_1 + \varphi_2 \geq 0$. Ist $\varphi_2 > 0$, dann erhalten wir analog $\varphi_1 + \varphi_2 > 0$. Dies zeigt die ersten beiden Aussagen. Sei nun $\lambda \geq 0$ und $\varphi \geq 0$. Dann ist auch $\lambda\varphi$ selbstadjungiert und wir erhalten $\langle v, (\lambda\varphi)(v) \rangle = \lambda\langle v, \varphi(v) \rangle \geq 0$, also $\lambda\varphi \geq 0$. Ist $\lambda > 0$ und $\varphi > 0$ dann erhalten wir mit dem gleichen Argument $\lambda\varphi > 0$. Behauptung (e) folgt aus $(\psi\varphi\psi^*)^* = \psi^{**}\varphi^*\psi^* = \psi\varphi\psi^*$ und $\langle (\psi\varphi\psi^*)(w), w \rangle = \langle \varphi(\psi^*(w)), \psi^*(w) \rangle \geq 0$, für alle $w \in W$. Für surjektives ψ ist ψ^* injektiv und wir erhalten wir daraus auch (f). Um auch (g) einzusehen, sei nun $\varphi \geq 0$ und $-\varphi \geq 0$, d.h. $\langle v, \varphi(v) \rangle = 0$, für alle $v \in V$. Mit der Polarisierungsidentität folgt $\langle w, \varphi(v) \rangle = 0$, für alle $v, w \in V$, und daher $\varphi = 0$. Ad (h): Ist $\varphi > 0$, dann auch $\ker(\varphi) = \{0\}$, also $\varphi: V \rightarrow V$ injektiv und daher invertierbar. \square

Positivität erlaubt es auf der Menge der selbstadjungierten Abbildungen eine Ordnungsrelation zu definieren. Sind $\varphi_1, \varphi_2: V \rightarrow V$ zwei selbstadjungierte Abbildungen dann schreiben wir $\varphi_2 \geq \varphi_1$ falls $\varphi_2 - \varphi_1 \geq 0$. Mit Hilfe des vorangehenden Lemmas erhalten wir sofort:

- (a) $\varphi \geq \varphi$.
- (b) Ist $\varphi_1 \geq \varphi_2$ und $\varphi_2 \geq \varphi_1$, dann gilt schon $\varphi_1 = \varphi_2$.

- (c) Ist $\varphi_1 \geq \varphi_2$ und $\varphi_2 \geq \varphi_3$, dann auch $\varphi_1 \geq \varphi_3$.
 (d) Ist $\varphi_1 \geq \tilde{\varphi}_1$ und $\varphi_2 \geq \tilde{\varphi}_2$, dann auch $\varphi_1 + \varphi_2 \geq \tilde{\varphi}_1 + \tilde{\varphi}_2$.
 (e) Ist $\lambda \geq 0$ ein Skalar und $\varphi \geq \psi$, dann auch $\lambda\varphi \geq \lambda\psi$.
 (f) Ist ψ linear und $\varphi_2 \geq \varphi_1$, dann auch $\psi\varphi_2\psi^* \geq \psi\varphi_1\psi^*$.

Dadurch wird die Menge der selbstadjungierten Abbildungen nur teilweise geordnet, i.A. sind zwei selbstadjungierte Abbildungen φ und ψ nicht vergleichbar, etwa gilt für $A = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$ weder $A \geq 0$ noch $0 \geq A$. Auch ist diese partielle Ordnung nicht mit der Multiplikation verträglich, d.h. aus $\varphi \geq 0$ und $\psi \geq 0$ lässt sich *nicht* $\varphi\psi \geq 0$ schließen, letztere Abbildung wird ja i.A. nicht einmal selbstadjungiert sein. Allerdings sind die semipositiven selbstadjungierten Abbildungen genau jene, die sich als Quadrate schreiben lassen. Genauer haben wir:

VII.5.6. SATZ. Für eine lineare Abbildung $\varphi: V \rightarrow V$ auf einem endlich dimensionalen Euklidischen oder unitären Vektorraum V sind äquivalent:

- (a) φ ist semipositiv.
 (b) φ ist selbstadjungiert und alle Eigenwerte sind größer oder gleich Null.
 (c) Es existiert eine Orthonormalbasis B von V , sodass $[\varphi]_{BB}$ eine Diagonalmatrix ist, deren Diagonaleinträge alle größer oder gleich Null sind.
 (d) Es existiert eine semipositive Abbildung $\rho: V \rightarrow V$, sodass $\rho^2 = \varphi$.
 (e) Es existiert eine selbstadjungierte Abbildung $\rho: V \rightarrow V$ mit $\rho^2 = \varphi$.
 (f) Es existiert eine lineare Abbildung $\rho: V \rightarrow V$ mit $\rho^*\rho = \varphi$.

Für $\varphi \geq 0$ ist die semipositive Abbildung $\rho^* = \rho \geq 0$ in (d) eindeutig bestimmt und wird mit $\sqrt{\varphi}$ oder $\varphi^{1/2}$ bezeichnet. Ist $\varphi \geq 0$ dann gilt:

- (g) Ist $\varphi = \lambda_1\pi_1 + \dots + \lambda_k\pi_k$ die Spektralzerlegung von φ , so ist die Spektralzerlegung von $\sqrt{\varphi}$ durch $\sqrt{\varphi} = \sqrt{\lambda_1}\pi_1 + \dots + \sqrt{\lambda_k}\pi_k$ gegeben. Insbesondere ist λ genau dann Eigenwert von φ wenn $\sqrt{\lambda}$ Eigenwert von $\sqrt{\varphi}$ ist und die entsprechenden Eigenräume stimmen überein.
 (h) $\text{img}(\sqrt{\varphi}) = \text{img}(\varphi)$
 (i) $\ker(\sqrt{\varphi}) = \ker(\varphi)$.
 (j) $\varphi > 0 \Leftrightarrow \sqrt{\varphi} > 0$.
 (k) Ist $\psi: V \rightarrow W$ eine Isometrie, so gilt $\sqrt{\psi\varphi\psi^*} = \psi\sqrt{\varphi}\psi^*$.
 (l) Ist $\psi: V \rightarrow V$ linear und $\psi\varphi = \varphi\psi$, dann auch $\psi\sqrt{\varphi} = \sqrt{\varphi}\psi$.

BEWEIS. Ad (a) \Rightarrow (b): Sei also $\varphi \geq 0$ und λ Eigenwert von φ . Es existiert daher $0 \neq v \in V$ mit $\varphi(v) = \lambda v$, wir erhalten

$$\lambda\|v\|^2 = \lambda\langle v, v \rangle = \langle v, \lambda v \rangle = \langle v, \varphi(v) \rangle \geq 0,$$

und somit $\lambda \geq 0$. Die Implikation (b) \Rightarrow (c) folgt aus dem Spektralsatz, siehe Satz VII.3.11. Ad (c) \Rightarrow (d): Sei also B eine Orthonormalbasis von V , sodass

$$[\varphi]_{BB} = \begin{pmatrix} a_1 & & \\ & \ddots & \\ & & a_n \end{pmatrix} \quad \text{und} \quad a_i \geq 0.$$

Definieren wir $\rho: V \rightarrow V$ durch

$$[\rho]_{BB} := \begin{pmatrix} \sqrt{a_1} & & \\ & \ddots & \\ & & \sqrt{a_n} \end{pmatrix},$$

dann gilt offensichtlich $\rho^* = \rho \geq 0$ und $\rho^2 = \varphi$. Die Implikationen (d) \Rightarrow (e) und (e) \Rightarrow (f) sind trivial. Nun zu (f) \Rightarrow (a): Jede Abbildung der Form $\varphi = \rho^* \rho$ ist selbstadjungiert, $\varphi^* = (\rho^* \rho)^* = \rho^* \rho^{**} = \rho^* \rho = \varphi$, und semipositiv, denn $\langle v, \varphi(v) \rangle = \langle v, (\rho^* \rho)(v) \rangle = \langle \rho(v), \rho(v) \rangle \geq 0$. Damit ist die Äquivalenz der Eigenschaften (a) bis (f) gezeigt.

Um die Eindeutigkeit der Wurzel zu zeigen, sei $\psi: V \rightarrow V$ mit $\psi \geq 0$ und $\psi^2 = \varphi$. Dann gilt $\psi\varphi = \psi\psi^2 = \psi^2\psi = \varphi\psi$, also bewahrt ψ jeden Eigenraum von φ . Bezeichne die Einschränkung von ψ auf einen solchen Eigenraum mit $\psi_i := \psi|_{E_{\lambda_i}^\varphi}: E_{\lambda_i}^\varphi \rightarrow E_{\lambda_i}^\varphi$. Beachte $\psi_i^* = \psi_i \geq 0$ und $\psi_i^2 = \lambda_i \text{id}_{E_{\lambda_i}^\varphi}$. Da ψ_i diagonalisierbar ist, folgt daraus $\psi_i = \sqrt{\lambda_i} \text{id}_{E_{\lambda_i}^\varphi}$, also $\psi|_{E_{\lambda_i}^\varphi} = \sqrt{\varphi}|_{E_{\lambda_i}^\varphi}$. Da dies für jeden Eigenraum gilt, folgt $\psi = \sqrt{\varphi}$, die Wurzel ist daher eindeutig.

Die verbleibenden Eigenschaften der Wurzel sind nun offensichtlich. \square

Wir definieren die Wurzel einer positiven Matrix mit Hilfe der damit assoziierten linearen Abbildung.

VII.5.7. BEISPIEL. Für Diagonalmatrizen mit positiven Einträgen gilt:

$$\begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}^{1/2} = \begin{pmatrix} \sqrt{\lambda_1} & & \\ & \ddots & \\ & & \sqrt{\lambda_n} \end{pmatrix}, \quad \lambda_i \geq 0.$$

VII.5.8. BEISPIEL. Wir wollen die Wurzel der positiven Matrix

$$A = \begin{pmatrix} 10 & 6 \\ 6 & 10 \end{pmatrix}$$

berechnen. Wir bestimmen zunächst eine orthogonale Matrix $U \in O_2$ und eine Diagonalmatrix D , sodass $A = UDU^{-1}$:

$$U = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}, \quad D = \begin{pmatrix} 16 & 0 \\ 0 & 4 \end{pmatrix}, \quad A = UDU^{-1} = UDU^t.$$

Es gilt daher

$$\sqrt{A} = \sqrt{UDU^{-1}} = U\sqrt{D}U^{-1} = U\sqrt{D}U^t = U \begin{pmatrix} 4 & 0 \\ 0 & 2 \end{pmatrix} U^t = \begin{pmatrix} 3 & 1 \\ 1 & 3 \end{pmatrix}.$$

VII.5.9. BEISPIEL. Wir wollen die Wurzel der positiven Matrix

$$A = \begin{pmatrix} 3 & 1 & -2 \\ 1 & 3 & -2 \\ -2 & -2 & 6 \end{pmatrix}$$

bestimmen. Wir berechnen zunächst eine orthogonale Matrix $U \in O_3$ und eine Diagonalmatrix D , sodass $A = UDU^{-1} = UDU^t$:

$$U = \begin{pmatrix} 1/\sqrt{3} & -1/\sqrt{2} & 1/\sqrt{6} \\ 1/\sqrt{3} & 1/\sqrt{2} & 1/\sqrt{6} \\ 1/\sqrt{3} & 0 & -2/\sqrt{6} \end{pmatrix}, \quad D = \begin{pmatrix} 2 & & \\ & 2 & \\ & & 8 \end{pmatrix}.$$

Für die Wurzel von A erhalten wir

$$\sqrt{A} = U\sqrt{D}U^t = U \begin{pmatrix} \sqrt{2} & & \\ & \sqrt{2} & \\ & & 2\sqrt{2} \end{pmatrix} U^t = \frac{\sqrt{2}}{6} \begin{pmatrix} 7 & 1 & -2 \\ 1 & 7 & -2 \\ -2 & -2 & 10 \end{pmatrix}.$$

VII.5.10. SATZ (Polarzerlegung). Sei $\varphi: V \rightarrow W$ eine lineare Abbildung zwischen endlich dimensionalen Euklidischen oder unitären Vektorräumen.

- (a) Ist $\dim(V) \leq \dim(W)$, dann lässt sich φ in der Form $\varphi = \psi\rho$ schreiben, wobei $\rho: V \rightarrow V$ semipositiv und $\psi: V \rightarrow W$ eine Isometrie ist. Dabei ist ρ eindeutig bestimmt, $\rho = \sqrt{\varphi^*\varphi}$. Ist φ injektiv, dann gilt $\rho > 0$ und $\psi = \varphi\rho^{-1}$, in diesem Fall ist also auch ψ eindeutig bestimmt.
- (b) Ist $\dim(V) \geq \dim(W)$, dann lässt sich φ in der Form $\varphi = \rho\psi^*$ schreiben, wobei $\rho: W \rightarrow W$ semipositiv und $\psi: W \rightarrow V$ eine Isometrie ist. Dabei ist ρ eindeutig bestimmt, $\rho = \sqrt{\varphi\varphi^*}$. Ist φ surjektiv, dann gilt $\rho > 0$ und $\psi = \varphi^*\rho^{-1}$, in diesem Fall ist also auch ψ eindeutig bestimmt.

BEWEIS. Ad (a): Ist $\varphi = \psi\rho$ eine Darstellung wie in (a), dann folgt

$$\varphi^*\varphi = (\psi\rho)^*(\psi\rho) = \rho^*\psi^*\psi\rho = \rho^*\rho = \rho^2,$$

also $\rho = \sqrt{\varphi^*\varphi}$. Ist φ injektiv, dann gilt $\ker(\rho) = \ker(\varphi^*\varphi) = \ker(\varphi) = 0$, also ist ρ invertierbar und $\psi = \varphi\rho^{-1}$. Damit sind die Eindeutigkeitsaussagen gezeigt.

Um die Existenz der Darstellung zu zeigen, sei zunächst $\varphi: V \rightarrow W$ injektiv. Dann ist $\rho := \sqrt{\varphi^*\varphi}$ invertierbar und wir definieren $\psi := \varphi\rho^{-1}$. Da ρ selbstadjungiert ist, gilt auch $(\rho^{-1})^* = \rho^{-1}$ und somit

$$\psi^*\psi = (\varphi\rho^{-1})^*(\varphi\rho^{-1}) = \rho^{-1}\varphi^*\varphi\rho^{-1} = \rho^{-1}\rho^2\rho^{-1} = \text{id}_V,$$

also ist ψ eine Isometrie. Damit ist die Existenz der gesuchten Zerlegung für injektive φ gezeigt.

Sei nun $\varphi: V \rightarrow W$ beliebig. Da die Einschränkung $\varphi|_{\ker(\varphi)^\perp}: \ker(\varphi)^\perp \rightarrow W$ injektiv ist, erhalten wir aus dem eben Bewiesenen eine (semi)positive Abbildung $\tilde{\rho}: \ker(\varphi)^\perp \rightarrow \ker(\varphi)^\perp$ und eine Isometrie $\tilde{\psi}: \ker(\varphi)^\perp \rightarrow W$, sodass

$$\tilde{\psi}\tilde{\rho} = \varphi|_{\ker(\varphi)^\perp}. \quad (\text{VII.26})$$

Aus bekannten Dimensionsformeln folgt

$$\begin{aligned} \dim \ker(\varphi) &= \dim V - \dim \text{img}(\varphi) \\ &= \dim V - \dim W + \dim \text{img}(\varphi)^\perp \leq \dim \text{img}(\varphi)^\perp, \end{aligned}$$

denn nach Voraussetzung ist $\dim V - \dim W \leq 0$. Es existiert daher eine Orthonormalbasis v_1, \dots, v_k von $\ker(\varphi)$ und ein Orthonormalsystem w_1, \dots, w_k von $\text{img}(\varphi)^\perp$. Die Abbildung $\tilde{\psi}: \ker(\varphi)^\perp \rightarrow W$ lässt sich daher zu einer Isometrie $\psi: V \rightarrow W$ fortsetzen, $\psi|_{\ker(\varphi)^\perp} = \tilde{\psi}$, $\psi(v_i) = w_i$, $i = 1, \dots, k$. Aus (VII.26) folgt:

$$\psi \iota \tilde{\rho} = \varphi|_{\ker(\varphi)^\perp},$$

wobei $\iota: \ker(\varphi)^\perp \rightarrow V$ die kanonische Inklusion bezeichnet. Daraus erhalten wir

$$\psi \iota \tilde{\rho} \pi = \varphi,$$

wobei $\pi: V \rightarrow \ker(\varphi)^\perp$ die Orthogonalprojektion bezeichnet. Es bleibt nur noch zu zeigen, dass die Abbildung $\rho := \iota \tilde{\rho} \pi: V \rightarrow V$ semipositiv ist. Da jedoch $\pi = \iota^*$, folgt dies aus Lemma VII.5.5(e).

Wir erhalten (b) indem wir (a) auf $\varphi^*: W \rightarrow V$ anwenden. \square

VII.5.11. KOROLLAR (Polarzerlegung für Matrizen).

a) Jede Matrix $A \in M_{n \times n}(\mathbb{C})$ lässt sich in der Form $A = UR$ schreiben, wobei $R^* = R \geq 0$ und $U \in U_n$. Dabei ist R eindeutig bestimmt, $R = \sqrt{A^*A}$. Für invertierbares A gilt $R > 0$ und $U = AR^{-1}$, in diesem Fall ist also auch U eindeutig bestimmt.

b) Jede Matrix $A \in M_{n \times n}(\mathbb{R})$ lässt sich in der Form $A = UR$ schreiben, wobei $R = R^t \geq 0$ und $U \in O_n$. Dabei ist R eindeutig bestimmt, $R = \sqrt{A^tA}$. Für invertierbares A gilt $R > 0$ und $U = AR^{-1}$, in diesem Fall ist also auch U eindeutig bestimmt.

VII.5.12. BEISPIEL. Ist $0 \neq a \in \mathbb{C}$, dann hat die Polarzerlegung der Matrix $A = (a) \in M_{1 \times 1}(\mathbb{C})$ die Form:

$$A = UR, \quad U = \left(\frac{a}{|a|}\right), \quad R = (|a|).$$

Die Polarzerlegung von Matrizen kann daher als Verallgemeinerung der klassischen Polardarstellung verstanden werden.

VII.5.13. BEISPIEL. Die Polarzerlegung einer Diagonalmatrix

$$A = \begin{pmatrix} a_1 & & \\ & \ddots & \\ & & a_n \end{pmatrix}, \quad a_i \in \mathbb{C},$$

ist $A = UR$, wobei

$$U = \begin{pmatrix} \frac{a_1}{|a_1|} & & \\ & \ddots & \\ & & \frac{a_n}{|a_n|} \end{pmatrix} \quad \text{und} \quad R = \begin{pmatrix} |a_1| & & \\ & \ddots & \\ & & |a_n| \end{pmatrix}.$$

Ist $a_j = 0$, dann ist $a_j/|a_j|$ nicht definiert und der entsprechende Diagonaleintrag in U kann beliebig (mit Absolutbetrag Eins) gewählt werden.

VII.5.14. BEISPIEL. Wir wollen die Polarzerlegung der invertierbaren Matrix

$$A = \begin{pmatrix} 13 & -15 \\ 9 & 5 \end{pmatrix}$$

bestimmen. Wir berechnen

$$A^*A = \begin{pmatrix} 13 & 9 \\ -15 & 5 \end{pmatrix} \begin{pmatrix} 13 & -15 \\ 9 & 5 \end{pmatrix} = 50 \begin{pmatrix} 5 & -3 \\ -3 & 5 \end{pmatrix}.$$

Durch Bestimmen einer Orthonormalbasis von Eigenvektoren der Matrix A^*A erhalten wir eine orthogonale Matrix $V \in O_2$ und eine Diagonalmatrix D , sodass:

$$A^*A = VDV^{-1} = VDV^t, \quad V = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}, \quad D = 100 \begin{pmatrix} 1 & \\ & 4 \end{pmatrix}.$$

Für die Wurzel folgt

$$\begin{aligned} R &= (A^*A)^{1/2} = VD^{1/2}V^t \\ &= \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} 10 \begin{pmatrix} 1 & \\ & 2 \end{pmatrix} \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} = 5 \begin{pmatrix} 3 & -1 \\ -1 & 3 \end{pmatrix}. \end{aligned}$$

Nach Konstruktion ist

$$U := AR^{-1} = \begin{pmatrix} 13 & -15 \\ 9 & 5 \end{pmatrix} \frac{1}{40} \begin{pmatrix} 3 & 1 \\ 1 & 3 \end{pmatrix} = \frac{1}{5} \begin{pmatrix} 3 & -4 \\ 4 & 3 \end{pmatrix}$$

eine orthogonale Matrix, und daher $A = UR$ die gesuchte Polarzerlegung von A .

VII.5.15. SATZ (Singulärwertzerlegung). Sei $\varphi: V \rightarrow W$ eine lineare Abbildung zwischen Euklidischen oder unitären Vektorräumen endlicher Dimension. Dann existieren eine Orthonormalbasis B von V , eine Orthonormalbasis C von W und $\sigma_1, \dots, \sigma_k > 0$, sodass

$$[\varphi]_{CB} = \begin{pmatrix} \sigma_1 & & 0 & \dots & 0 \\ & \ddots & \vdots & & \vdots \\ 0 & \dots & \sigma_k & 0 & \dots & 0 \\ \vdots & & \vdots & \vdots & & \vdots \\ 0 & \dots & 0 & 0 & \dots & 0 \end{pmatrix}. \quad (\text{VII.27})$$

Dabei sind $\sigma_1, \dots, \sigma_k$, die sogenannten Singulärwerte von φ , bis auf ihre Reihenfolge eindeutig bestimmt. Es gilt nämlich $k = \text{rank}(\varphi^*\varphi) = \text{rank}(\varphi\varphi^*)$ und $\sigma_1^2, \dots, \sigma_k^2$ sind die nicht-trivialen Eigenwerte von $\varphi^*\varphi$ und gleichzeitig auch die nicht-trivialen Eigenwerte von $\varphi\varphi^*$.

Darüber hinaus ist $\sigma > 0$ genau dann Singulärwert von φ , wenn Einheitsvektoren $v \in V$ und $w \in W$ existieren, $\|v\| = 1 = \|w\|$, sodass $\varphi(v) = \sigma w$ und $\varphi^*(w) = \sigma v$.

BEWEIS. Da $\varphi^*\varphi \geq 0$, existiert eine Orthonormalbasis $B = (b_1, \dots, b_n)$ aus Eigenvektoren, d.h. $(\varphi^*\varphi)(b_i) = \lambda_i b_i$, und für die Eigenwerte gilt $\lambda_i \geq 0$. Durch

Umbenennen der Basiselemente können wir daher $\lambda_1, \dots, \lambda_k > 0$ und $\lambda_{k+1} = \dots = \lambda_n = 0$ erreichen. Wir setzen

$$\sigma_i := \sqrt{\lambda_i} \quad \text{und} \quad c_i := \frac{1}{\sigma_i} \varphi(b_i), \quad i = 1, \dots, k.$$

Beachte, dass die Vektoren c_1, \dots, c_k ein Orthonormalsystem in W bilden, denn

$$\begin{aligned} \langle c_i, c_j \rangle &= \frac{1}{\sigma_i \sigma_j} \langle \varphi(b_i), \varphi(b_j) \rangle = \frac{1}{\sigma_i \sigma_j} \langle b_i, (\varphi^* \varphi)(b_j) \rangle \\ &= \frac{1}{\sigma_i \sigma_j} \langle b_i, \lambda_j b_j \rangle = \frac{\lambda_j}{\sigma_i \sigma_j} \langle b_i, b_j \rangle = \frac{\lambda_j}{\sigma_i \sigma_j} \delta_{ij} = \delta_{ij}. \end{aligned}$$

Dieses lässt sich zu einer Orthonormalbasis, $C = (c_1, \dots, c_k, c_{k+1}, \dots, c_m)$ von W ergänzen. Nach Konstruktion hat die Matrixdarstellung von φ die Form (VII.27). Die restlichen Behauptungen sind nun offensichtlich. \square

Für Matrizen erhalten wir daraus:

VII.5.16. KOROLLAR (Singularwertzerlegung für Matrizen).

a) Jede reelle Matrix $A \in M_{m \times n}(\mathbb{R})$ lässt sich in der Form $A = U \Sigma V^t$ schreiben, wobei $U \in O_m$, $V \in O_n$ und Σ wie in (VII.27).

b) Jede komplexe Matrix $A \in M_{m \times n}(\mathbb{C})$ lässt sich in der Form $A = U \Sigma V^*$ schreiben, wobei $U \in U_m$, $V \in U_n$ und Σ wie in (VII.27).

BEWEIS. Wir zeigen nur b), der reelle Fall lässt sich analog behandeln. Es bezeichne $\psi: \mathbb{C}^n \rightarrow \mathbb{C}^m$, $\psi(x) = Ax$, die mit A assoziierte lineare Abbildung. Nach Satz VII.5.15 existieren eine Orthonormalbasis $B = (b_1, \dots, b_n)$ von \mathbb{C}^n und eine Orthonormalbasis $C = (c_1, \dots, c_m)$ von \mathbb{C}^m , sodass die Matrixdarstellung $\Sigma := [\psi]_{CB}$ die Gestalt (VII.27) hat. Da B und C Orthonormalbasen bilden, sind die Basiswechselmatrizen $V := T_{EB} = (b_1 | \dots | b_n)$ und $U := T_{\tilde{E}C} = (c_1 | \dots | c_m)$ unitär, wobei E die Standardbasis von \mathbb{C}^n und \tilde{E} die Standardbasis von \mathbb{C}^m bezeichnen. Es gilt daher $A = [\psi]_{\tilde{E}E} = T_{\tilde{E}C} [\psi]_{CB} T_{EB}^{-1} = U \Sigma V^{-1} = U \Sigma V^*$. \square

Beachte, dass die orthogonalen/unitären Matrizen U und V i.A. nicht eindeutig bestimmt sind.

VII.5.17. BEISPIEL. Wir wollen eine Singularwertzerlegung der Matrix

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

bestimmen. Wir gehen genau wie im Beweis des vorangehenden Satzes vor und berechnen zunächst:

$$A^* A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 2 \\ 0 & 2 & 2 \end{pmatrix}.$$

Es ist dann

$$b_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad b_2 = \begin{pmatrix} 0 \\ 1/\sqrt{2} \\ 1/\sqrt{2} \end{pmatrix}, \quad b_3 = \begin{pmatrix} 0 \\ 1/\sqrt{2} \\ -1/\sqrt{2} \end{pmatrix}$$

eine Orthonormalbasis aus Eigenvektoren von A^*A mit Eigenwerten 1, 4, 0. Somit

$$\sigma_1 = 1, \quad \sigma_2 = 2, \quad c_1 = \frac{1}{\sigma_1}Ab_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \quad c_2 = \frac{1}{\sigma_2}Ab_2 = \begin{pmatrix} 0 \\ 1/\sqrt{2} \\ -1/\sqrt{2} \\ 0 \end{pmatrix}.$$

Wir ergänzen dies durch

$$c_3 = \begin{pmatrix} 0 \\ 1/\sqrt{2} \\ 1/\sqrt{2} \\ 0 \end{pmatrix}, \quad c_4 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}$$

zu einer Orthonormalbasis und erhalten die Singulärwertzerlegung

$$A = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1/\sqrt{2} & 1/\sqrt{2} & 0 \\ 0 & -1/\sqrt{2} & 1/\sqrt{2} & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1/\sqrt{2} & 1/\sqrt{2} \\ 0 & 1/\sqrt{2} & -1/\sqrt{2} \end{pmatrix}^*.$$

Wir wollen diesen Abschnitt mit einigen Bemerkungen zu anderen Funktionen von Matrizen beenden. Wir erläutern dies an einem der wichtigsten Beispiele, der *Exponentialfunktion*. Für jede quadratische Matrix A setzen wir:

$$\exp(A) = e^A = \sum_{n=0}^{\infty} \frac{1}{n!} A^n. \quad (\text{VII.28})$$

Mit Hilfe der weiter oben besprochenen Operatornorm erhalten wir

$$\sum_{n=0}^{\infty} \frac{1}{n!} \|A^n\| \leq \sum_{n=0}^{\infty} \frac{1}{n!} \|A\|^n = e^{\|A\|} < \infty$$

also konvergiert die Reihe (VII.28) für jede quadratische Matrix A und es gilt

$$\|e^A\| \leq e^{\|A\|}.$$

Etwa gilt für Diagonalmatrizen:

$$\exp \begin{pmatrix} a_1 & & \\ & \ddots & \\ & & a_n \end{pmatrix} = \begin{pmatrix} e^{a_1} & & \\ & \ddots & \\ & & e^{a_n} \end{pmatrix}.$$

Für jede invertierbare Matrix U haben wir:

$$e^{UAU^{-1}} = Ue^AU^{-1}. \quad (\text{VII.29})$$

I.A. ist $e^{AB} \neq e^A e^B$, für kommutierende A und B gilt jedoch Gleichheit. Insbesondere ist $t \mapsto e^{tA}$ ein Gruppenhomomorphismus:

$$e^{(s+t)A} = e^{sA} e^{tA}, \quad (e^A)^{-1} = e^{-A} \quad \text{und} \quad \frac{d}{dt} e^{tA} = A e^{tA}.$$

Für jedes v_0 ist daher $v(t) = e^{tA} v_0$ die Lösung des Systems gewöhnlicher Differentialgleichung $v'(t) = Av(t)$ zum Anfangswert $v(0) = v_0$.

VII.5.18. BEISPIEL. Wir wollen e^{tA} bestimmen, wobei

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

Durch Berechnung einer Eigenbasis von A erhalten wir die Darstellung:

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

Mit (VII.29) daher

$$\begin{aligned} e^{tA} &= \begin{pmatrix} 1 & 1 \\ -\mathbf{i} & \mathbf{i} \end{pmatrix} \begin{pmatrix} e^{t\mathbf{i}} & 0 \\ 0 & e^{-t\mathbf{i}} \end{pmatrix} \begin{pmatrix} 1 & 1 \\ -\mathbf{i} & \mathbf{i} \end{pmatrix}^{-1} \\ &= \begin{pmatrix} \frac{e^{t\mathbf{i}} + e^{-t\mathbf{i}}}{2} & -\frac{e^{t\mathbf{i}} - e^{-t\mathbf{i}}}{2\mathbf{i}} \\ \frac{e^{t\mathbf{i}} - e^{-t\mathbf{i}}}{2\mathbf{i}} & \frac{e^{t\mathbf{i}} + e^{-t\mathbf{i}}}{2} \end{pmatrix} = \begin{pmatrix} \cos t & -\sin t \\ \sin t & \cos t \end{pmatrix}, \end{aligned}$$

d.h. e^{tA} ist eine Drehung um den Winkel t .

Analog lassen sich konvergente Potenzreihe, $f(z) = \sum_n a_n z^n$, auf quadratische Matrizen anwenden, $f(A) := \sum_n a_n A^n$. So lässt sich etwa der Gleichung

$$e^{\mathbf{i}A} = \cos(A) + \mathbf{i} \sin(A),$$

auch für quadratische Matrizen A Sinn geben. Basisunabhängig, d.h. für Endomorphismen φ eines endlich dimensionalen Vektorraums: $f(\varphi) = \sum_n a_n \varphi^n$.

VII.6. Moore–Penrose Pseudoinverse. Sei $\varphi: V \rightarrow W$ eine lineare Abbildung zwischen endlich dimensionalen Euklidischen oder unitären Vektorräumen. Dann ist die Einschränkung,

$$\varphi|_{\ker(\varphi)^\perp} : \ker(\varphi)^\perp \xrightarrow{\cong} \text{img}(\varphi)$$

invertierbar und wir definieren ihre sogenannte *Moore–Penrose Pseudoinverse*, $\varphi^+ : W \rightarrow V$, durch

$$\varphi^+|_{\text{img}(\varphi)} := \varphi|_{\ker(\varphi)^\perp}^{-1} \quad \text{und} \quad \varphi^+|_{\text{img}(\varphi)^\perp} := 0.$$

VII.6.1. SATZ (Moore–Penrose Pseudoinverse). *Sei $\varphi: V \rightarrow W$ eine lineare Abbildung zwischen endlich dimensionalen Euklidischen oder unitären Vektorräumen. Dann hat die Pseudoinverse, $\varphi^+ : W \rightarrow V$, folgende vier Eigenschaften und ist dadurch eindeutig charakterisiert:*

(a) $\varphi\varphi^+\varphi = \varphi$.

- (b) $\varphi^+\varphi\varphi^+ = \varphi^+$.
 (c) $\varphi\varphi^+$ ist selbstadjungiert, d.h. $(\varphi\varphi^+)^* = \varphi\varphi^+$.
 (d) $\varphi^+\varphi$ ist selbstadjungiert, d.h. $(\varphi^+\varphi)^* = \varphi^+\varphi$.

Darüber hinaus gilt:

- (e) $\varphi^*\varphi\varphi^+ = \varphi^*$.
 (f) $\varphi^+\varphi\varphi^* = \varphi^*$.
 (g) $\varphi\varphi^+$ ist die Orthogonalprojektion auf $\text{img}(\varphi) = \ker(\varphi^*)^\perp$.
 (h) $\varphi^+\varphi$ ist die Orthogonalprojektion auf $\text{img}(\varphi^*) = \ker(\varphi)^\perp$.
 (i) $\ker(\varphi^+) = \ker(\varphi^*) = \text{img}(\varphi)^\perp$.
 (j) $\text{img}(\varphi^+) = \text{img}(\varphi^*) = \ker(\varphi)^\perp$.
 (k) Ist φ invertierbar, dann gilt $\varphi^+ = \varphi^{-1}$.
 (l) Ist φ surjektiv, so ist $\varphi\varphi^+ = \text{id}_W$ und $\varphi^+ = \varphi^*(\varphi\varphi^*)^{-1}$.
 (m) Ist φ injektiv, so ist $\varphi^+\varphi = \text{id}_V$ und $\varphi^+ = (\varphi^*\varphi)^{-1}\varphi^*$.
 (n) Ist $\psi: U \rightarrow V$ surjektiv und $\varphi: V \rightarrow W$ injektiv, dann gilt $(\varphi\psi)^+ = \psi^+\varphi^+ = \psi^*(\psi\psi^*)^{-1}(\varphi^*\varphi)^{-1}\varphi^* = \psi^*(\varphi^*\varphi\psi\psi^*)^{-1}\varphi^*$.
 (o) $(\lambda\varphi)^+ = \lambda^{-1}\varphi^+$, für alle $0 \neq \lambda \in \mathbb{K}$.
 (p) $(\varphi^*)^+ = (\varphi^+)^*$.
 (q) Sei b_1, \dots, b_k eine Basis von $\ker(\varphi)^\perp$ und b_{k+1}, \dots, b_n eine Basis von $\ker(\varphi)$. Weiters sei c_1, \dots, c_k eine Basis von $\text{img}(\varphi)$ und c_{k+1}, \dots, c_m eine Basis von $\text{img}(\varphi)^\perp$. Dann gilt

$$[\varphi]_{CB} = \begin{pmatrix} A & 0 \\ 0 & 0 \end{pmatrix} \quad \text{und} \quad [\varphi^+]_{BC} = \begin{pmatrix} A^{-1} & 0 \\ 0 & 0 \end{pmatrix}.$$

wobei A eine invertierbare $(k \times k)$ -Matrix bezeichnet.

BEWEIS. Nach Konstruktion ist $\varphi\varphi^+|_{\text{img}(\varphi)} = \text{id}_{\text{img}(\varphi)}$ und $\varphi\varphi^+|_{\text{img}(\varphi)^\perp} = 0$, also stimmt $\varphi\varphi^+$ mit der Orthogonalprojektion auf $\text{img}(\varphi)$ überein. Dies zeigt (a), (c) und (g). Nach Konstruktion gilt $\varphi^+\varphi|_{\ker(\varphi)^\perp} = \text{id}_{\ker(\varphi)^\perp}$ und offensichtlich auch $\varphi^+\varphi|_{\ker(\varphi)} = 0$. Daraus schließen wir, dass $\varphi^+\varphi$ mit der Orthogonalprojektion auf $\ker(\varphi)^\perp$ übereinstimmt. Dies zeigt (b), (d) und (h). Behauptung (e) folgt aus (a) und (c). Analog erhalten wir (f) sofort aus (a) und (d). Auch (i), (j), (k), (o) und (q) folgen sofort aus der Definition. Behauptung (l) folgt aus (g) und (f), denn für surjektives φ ist $\varphi\varphi^*$ invertierbar. Analog erhalten wir aus (h) und (e) sofort (m), denn für injektives φ ist $\varphi^*\varphi$ invertierbar. Ad (n): In diesem Fall gilt $\ker(\varphi\psi) = \ker(\psi)$ und $\text{img}(\varphi\psi) = \text{img}(\varphi)$ und aus unserer Definition folgt sofort $(\varphi\psi)^+ = \psi^+\varphi^+$. Die anderen Formeln für $\psi^+\varphi^+$ folgen dann aus (l) und (m).

Wir verifizieren nun, dass die Pseudoinverse durch die Eigenschaften (a) bis (d) eindeutig bestimmt ist. Sei dazu $\tilde{\varphi}^+: W \rightarrow V$, sodass (a) bis (d) auch für $\tilde{\varphi}^+$ statt φ^+ gelten. Setzen wir $p := \tilde{\varphi}^+\varphi$, so folgt $p^2 = p = p^*$ und $\ker(p) = \ker(\varphi)$ aus (a) und (d). Somit ist p die Orthogonalprojektion auf $\ker(\varphi)$, also

$$\tilde{\varphi}^+\varphi = p = \varphi^+\varphi.$$

Analog erhalten wir für $q := \varphi\tilde{\varphi}^+$ aus (a) und (c) sofort $q^2 = q = q^*$ und $\text{img}(q) = \text{img}(\varphi)$, also ist q die Orthogonalprojektion auf $\text{img}(\varphi)$ und daher

$$\varphi\tilde{\varphi}^+ = q = \varphi\varphi^+.$$

Mit (d) folgt

$$\tilde{\varphi}^+ = \tilde{\varphi}^+\varphi\tilde{\varphi}^+ = \varphi^+\varphi\tilde{\varphi}^+ = \varphi^+\varphi\varphi^+ = \varphi^+,$$

womit die Eindeutigkeitsaussage gezeigt wäre.

Ad (p): Aus (a) bis (d) erhalten wir $\varphi^*(\varphi^+)^*\varphi^* = \varphi^*$, $(\varphi^+)^*\varphi^*(\varphi^+)^* = (\varphi^+)^*$, und $\varphi^*(\varphi^+)^*$ sowie $(\varphi^+)^*\varphi^*$ sind selbstadjungiert. Aus der eben bewiesenen Eindeutigkeit der Pseudoinversen folgt daher $(\varphi^*)^+ = (\varphi^+)^*$. \square

Für Matrizen erhalten wir daraus:

VII.6.2. KOROLLAR. Sei $\mathbb{K} = \mathbb{R}$ oder $\mathbb{K} = \mathbb{C}$. Für jede Matrix $A \in M_{m \times n}(\mathbb{K})$ existiert eine eindeutig bestimmte Matrix $A^+ \in M_{n \times m}(\mathbb{K})$ mit folgenden vier Eigenschaften:

- (a) $AA^+A = A$
- (b) $A^+AA^+ = A^+$
- (c) $(AA^+)^* = AA^+$
- (d) $(A^+A)^* = A^+A$

Diese Pseudoinverse A^+ hat die zu den in Satz VII.6.1 analogen Eigenschaften.

- (e) $A^*AA^+ = A^*$
- (f) $A^+AA^* = A^*$
- (g) AA^+ ist die Matrix der Orthogonalprojektion auf $\text{img}(A) = \ker(A^*)^\perp$.
- (h) A^+A ist die Matrix der Orthogonalprojektion auf $\text{img}(A^*) = \ker(A)^\perp$.
- (i) $\ker(A^+) = \ker(A^*) = \text{img}(A)^\perp$.
- (j) $\text{img}(A^+) = \text{img}(A^*) = \ker(A)^\perp$.
- (k) Ist A invertierbar, dann gilt $A^+ = A^{-1}$.
- (l) Ist $\text{rank}(A) = m$, so gilt $AA^+ = I_m$ und $A^+ = A^*(AA^*)^{-1}$.
- (m) Ist $\text{rank}(A) = n$, so gilt $A^+A = I_n$ und $A^+ = (A^*A)^{-1}A^*$.
- (n) Ist $B \in M_{n \times l}(\mathbb{K})$ und $\text{rank}(A) = n = \text{rank}(B)$, dann gilt $(AB)^+ = B^+A^+ = B^*(BB^*)^{-1}(A^*A)^{-1}A^* = B^*(A^*ABB^*)^{-1}A^*$.
- (o) $(\lambda A)^+ = \lambda^{-1}A^+$, für alle Skalare $0 \neq \lambda \in \mathbb{K}$.
- (p) $(A^*)^+ = (A^+)^*$

VII.6.3. BEISPIEL. Ist $A \in \text{GL}_k(\mathbb{K})$, dann gilt:

$$\underbrace{\begin{pmatrix} A & 0 \\ 0 & 0 \end{pmatrix}}_{\in M_{m \times n}(\mathbb{K})}^+ = \underbrace{\begin{pmatrix} A^{-1} & 0 \\ 0 & 0 \end{pmatrix}}_{\in M_{n \times m}(\mathbb{K})}.$$

VII.6.4. BEISPIEL. Nach Korollar VII.6.2(1) gilt:

$$\begin{aligned} \underbrace{\begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \end{pmatrix}}_A^+ &= A^*(AA^*)^{-1} = \begin{pmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{pmatrix} \left\{ \begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{pmatrix} \right\}^{-1} \\ &= \begin{pmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{pmatrix} \begin{pmatrix} 3 & 6 \\ 6 & 14 \end{pmatrix}^{-1} = \begin{pmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{pmatrix} \frac{1}{6} \begin{pmatrix} 14 & -6 \\ -6 & 3 \end{pmatrix} = \begin{pmatrix} 4/3 & -1/2 \\ 1/3 & 0 \\ -2/3 & 1/2 \end{pmatrix}. \end{aligned}$$

VII.6.5. BEISPIEL. Wir wollen die Pseudoinverse von

$$A = \begin{pmatrix} 1 & 0 & 1 & 0 \\ 1 & 2 & 3 & 4 \\ 2 & 2 & 4 & 4 \end{pmatrix}$$

bestimmen. Zunächst mittels Zeilenumformungen:

$$\begin{aligned} A &= \begin{pmatrix} 1 & 0 & 1 & 0 \\ 1 & 2 & 3 & 4 \\ 2 & 2 & 4 & 4 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ -1 & -1 & 1 \end{pmatrix}^{-1} \begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 2 & 2 & 4 \\ 0 & 0 & 0 & 0 \end{pmatrix} \\ &= \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 2 & 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 2 & 2 & 4 \\ 0 & 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 1 & 1 \\ 2 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 2 & 2 & 4 \end{pmatrix} =: LU \end{aligned}$$

Mit Korollar VII.6.2(n) daher:

$$\begin{aligned} A^+ &= U^+L^+ = U^*(L^*LUU^*)^{-1}L^* \\ &= U^* \left\{ \begin{pmatrix} 6 & 3 \\ 3 & 2 \end{pmatrix} \begin{pmatrix} 2 & 2 \\ 2 & 24 \end{pmatrix} \right\}^{-1} L^* = U^* \begin{pmatrix} 18 & 84 \\ 10 & 54 \end{pmatrix}^{-1} L^* \\ &= \begin{pmatrix} 1 & 0 \\ 0 & 2 \\ 1 & 2 \\ 0 & 4 \end{pmatrix} \frac{1}{66} \begin{pmatrix} 27 & -42 \\ -5 & 9 \end{pmatrix} \begin{pmatrix} 1 & 1 & 2 \\ 0 & 1 & 1 \end{pmatrix} = \frac{1}{66} \begin{pmatrix} 27 & -15 & 12 \\ -10 & 8 & -2 \\ 17 & -7 & 10 \\ -20 & 16 & -4 \end{pmatrix}. \end{aligned}$$

Als Anwendung der Pseudoinversen wollen wir nun lineare Gleichungssysteme, $Ax = y$, betrachten, die sich nicht exakt lösen lassen. In solch einer Situation ist es oft hilfreich x so zu bestimmen, dass der Euklidische Abstand, $d(Ax, y) = \|Ax - y\|$, minimal wird.

VII.6.6. SATZ (Methode der kleinsten Quadrate). Sei $\mathbb{K} = \mathbb{R}$ oder $\mathbb{K} = \mathbb{C}$. Weiters seien $A \in M_{m \times n}(\mathbb{K})$, $y \in \mathbb{K}^m$ und $x \in \mathbb{K}^n$. Dann sind äquivalent:

- $d(Ax, y) \leq d(A\tilde{x}, y)$, für alle $\tilde{x} \in \mathbb{K}^n$. (x liefert beste Approximation)
- $Ax - y \perp \text{img}(A)$.
- $A^*Ax = A^*y$. (Normalgleichungen)

(d) $x \in A^+y + \ker(A)$, wobei $A^+ \in M_{n \times m}(\mathbb{K})$ die Pseudoinverse von A bezeichnet. Unter allen x , die diesen Bedingungen genügen, ist $x = A^+y$ das eindeutige mit kleinster Norm.

BEWEIS. Die Äquivalenz (a) \Leftrightarrow (b) folgt aus Satz VII.2.32. Die Äquivalenz (b) \Leftrightarrow (c) folgt aus $\text{img}(A)^\perp = \ker(A^*)$, siehe Proposition VII.3.2. Ad (d) \Leftrightarrow (c): Ein Vektor $x \in \mathbb{R}^n$ genügt genau dann den Normalgleichungen in (c), wenn

$$A^*AA^+y + A^*A(x - A^+y) = A^*y.$$

Nach Korollar VII.6.2(e) gilt $A^*AA^+ = A^*$, also ist dies zu $A^*A(x - A^+y) = 0$ äquivalent. Dies ist genau dann der Fall, wenn $x - A^+y \in \ker(A^*A) = \ker(A)$ gilt, d.h. genau dann wenn $x \in A^+y + \ker(A)$. Damit ist die Äquivalenz von (a) bis (d) gezeigt. Ist $\xi \in \ker(A)$, dann gilt $\xi \perp A^+y$, siehe Korollar VII.6.2(j), und nach dem Satz von Pythagoras daher

$$\|A^+y + \xi\|^2 = \|A^+y\|^2 + \|\xi\|^2 \geq \|A^+y\|^2,$$

wobei Gleichheit nur für $\xi = 0$ gilt. Dies zeigt, dass $x = A^+y$ die eindeutige Lösung der Normalgleichungen mit kleinster Norm ist. \square

VII.6.7. BEISPIEL (Polynomausgleich). Gegeben seien $x_1, \dots, x_N \in \mathbb{R}$ und $y_1, \dots, y_N \in \mathbb{R}$, sowie $n \in \mathbb{N}$. Wir suchen ein Polynom n -ten Grades,

$$p(x) = a_0 + a_1x + a_2x^2 + \dots + a_nx^n,$$

sodass $p(x_i)$ möglichst nahe bei y_i liegt. Genauer, die Summe der Fehlerquadrate,

$$\sum_{i=1}^N |p(x_i) - y_i|^2,$$

soll minimal sein. Betrachten wir

$$A = \begin{pmatrix} 1 & x_0 & x_0^2 & \cdots & x_0^n \\ 1 & x_1 & x_1^2 & \cdots & x_1^n \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_N & x_N^2 & \cdots & x_N^n \end{pmatrix}, \quad a = \begin{pmatrix} a_0 \\ a_1 \\ \vdots \\ a_n \end{pmatrix} \quad \text{und} \quad y = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_N \end{pmatrix},$$

dann gilt:

$$Aa = \begin{pmatrix} p(x_0) \\ p(x_1) \\ \vdots \\ p(x_N) \end{pmatrix} \quad \text{und} \quad d(Aa, y)^2 = \sum_{i=1}^N |p(x_i) - y_i|^2$$

Es ist daher $a \in \mathbb{R}^{n+1}$ so zu bestimmen, sodass $d(Aa, y)$ minimal wird. Nach Satz VII.6.6 ist dies für $a = A^+y$ der Fall. Haben wir wenigstens $n + 1$ verschiedene Werte x_i vorliegen, dann hat A maximalen Rang, $\text{rank}(A) = n + 1$, siehe

Satz IV.6.26. Nach Korollar VII.6.2(m) gilt in diesem Fall $A^+ = (A^*A)^{-1}A^*$, die Koeffizienten des gesuchten Polynoms sind daher durch

$$a = (A^*A)^{-1}A^*y$$

gegeben.

VII.6.8. BEISPIEL (Linearer Ausgleich). Es seien x_i und y_i wie folgt gegeben:

i	1	2	3	4	5	6	7
x_i	2.5	3.1	3.1	4.2	5.1	6.7	6.7
y_i	6.5	7.5	7.9	9.9	11.8	14.9	14.5

Gesucht sei ein lineares Polynom, $p(x) = a + bx$, sodass die Summe der Fehlerquadrate,

$$\sum_{i=1}^7 |p(x_i) - y_i|^2$$

minimal wird. Nach Beispiel VII.6.7 sind die Koeffizienten durch

$$\begin{pmatrix} a \\ b \end{pmatrix} = (A^*A)^{-1}A^*y = \frac{1}{126.34} \begin{pmatrix} 209.978 \\ 246.91 \end{pmatrix} = \begin{pmatrix} 1.6620073\dots \\ 1.9543296\dots \end{pmatrix}$$

gegeben, wobei

$$A = \begin{pmatrix} 1 & 2.5 \\ 1 & 3.1 \\ 1 & 3.1 \\ 1 & 4.2 \\ 1 & 5.1 \\ 1 & 6.7 \\ 1 & 6.7 \end{pmatrix} \quad \text{und} \quad y = \begin{pmatrix} 6.5 \\ 7.5 \\ 7.9 \\ 9.9 \\ 11.8 \\ 14.9 \\ 14.5 \end{pmatrix}.$$

Denn

$$(A^*A)^{-1} = \begin{pmatrix} 7 & 31.4 \\ 31.4 & 158.9 \end{pmatrix}^{-1} = \frac{1}{126.34} \begin{pmatrix} 158.9 & -31.4 \\ -31.4 & 7 \end{pmatrix},$$

und

$$(A^*A)^{-1}A^* = \frac{1}{126.34} \begin{pmatrix} 80.4 & 61.56 & 61.56 & 27.02 & -1.24 & -51.48 & -51.48 \\ -13.9 & -9.7 & -9.7 & -2 & 4.3 & 15.5 & 15.5 \end{pmatrix}.$$

Das gesuchte lineare Polynom ist daher:

$$p(x) = a + bx \sim 1.66 + 1.95x.$$

VIII. Affine und projektive Geometrie

Wir beginnen dieses Kapitel mit einem Abschnitt, in dem wir grundlegende Konzepte der affinen Geometrie besprechen: affine Abbildungen, affine Teilräume, affine Hüllen, affine Unabhängigkeit, affine Koordinatensysteme. Alles lässt sich hier sehr leicht auf den linearen Fall zurückführen. Wir vermeiden das Konzept des affinen Raums über einem Vektorraum [11] und betrachten nur affine Abbildungen zwischen Vektorräumen, bzw. affine Teilräume von Vektorräumen. Am Ende des ersten Abschnitts geben wir einen Beweis des Fundamentalsatzes der affinen Geometrie, siehe Satz VIII.1.32. Dieser besagt, dass affine Isomorphismen als jene Bijektionen charakterisiert werden können, die Geraden auf Geraden abbilden.

Im zweiten Abschnitt betrachten wir dann Quadriken, d.h. Teilmengen, die sich durch eine quadratische Gleichung beschreiben lassen. Das erste Resultat ist eine Klassifikation aller Quadriken, bis auf affine Isomorphismen. Wir werden zeigen, dass jede Quadrik in \mathbb{C}^n bzw. \mathbb{R}^n zu einer von endlich vielen, sehr expliziten Normalformen affin kongruent ist. Dies ist eine grobe Klassifikation, etwa kann jede Ellipse durch einen affinen Isomorphismus auf den Einheitskreis abgebildet werden und ist daher zum Einheitskreis affin kongruent. Neben den Ellipsen treffen wir hier auch Hyperbeln, Parabeln und einige weitere, degenerierte Fälle an. Dabei ist die komplexe Klassifikation, siehe Korollar VIII.2.10, etwas einfacher als die reelle, siehe Korollar VIII.2.13. Dies hängt damit zusammen, dass die Klassifikation der komplexen quadratischen Form (Rang) einfacher ist als die Klassifikation der reellen quadratischen Formen (Signatur). Anschließend diskutieren wir die metrische Klassifikation reeller Quadriken. Wir werden sehen, dass jede reelle Quadrik durch eine Bewegung auf eine Normalform in Hauptlage gebracht werden kann. Die Liste der Normalformen, siehe Korollar VIII.2.19, hat nun auch kontinuierliche Parameter, etwa können zwei Ellipsen nur dann metrisch kongruent sein, wenn sie die gleichen Halbachsen haben.

Nachdem wir im dritten Abschnitt projektive Räume und Abbildungen besprochen haben wenden wir uns im vierten Abschnitt dann den Quadriken in projektiven Räumen zu. Die Klassifikation der projektiven Quadriken ist einfacher als die Klassifikation der affinen Quadriken. Etwa sind Ellipsen, Hyperbeln und Parabeln alle zum Kreis projektiv kongruent: wir können eine Parabel als eine Kreis verstehen, der die Ferngerade in einem Punkt schneidet; und eine Hyperbel ist ein Kreis, der die Ferngerade in zwei Punkten schneidet.

Wir orientieren uns hier an der Darstellung [11], siehe aber auch [12].

VIII.1. Affine Räume und Abbildungen. Sei V ein Vektorraum über \mathbb{K} . Unter einer *Affinkombination* von $v_1, \dots, v_k \in V$ verstehen wir jeden Ausdruck der Form

$$\lambda_1 v_1 + \dots + \lambda_k v_k,$$

wobei $\lambda_1, \dots, \lambda_k \in \mathbb{K}$ und $\sum_{i=1}^k \lambda_i = 1$.

VIII.1.1. BEISPIEL. Sind v_1 und v_2 zwei verschiedene Punkte in \mathbb{R}^3 , dann stimmt die Menge der Affinkombinationen, $\{\lambda_1 v_1 + \lambda_2 v_2 \mid \lambda_1 + \lambda_2 = 1\}$, mit der Geraden durch v_1 und v_2 überein. Sind $v_1, v_2, v_3 \in \mathbb{R}^3$ drei Punkte, die nicht auf einer Geraden liegen, dann ist $\{\sum_{i=1}^3 \lambda_i v_i \mid \sum_{i=1}^3 \lambda_i = 1\}$ die von v_1, v_2 und v_3 aufgespannte Ebene.

VIII.1.2. DEFINITION (Affine Abbildung). Seien V und W zwei Vektorräume über \mathbb{K} . Eine Abbildung $\alpha: V \rightarrow W$ wird *affin* genannt wenn sie mit Affinkombinationen verträglich ist, d.h. wenn für alle $v_1, \dots, v_k \in V$ und alle $\lambda_1, \dots, \lambda_k \in \mathbb{K}$ mit $\sum_{i=1}^k \lambda_i = 1$ gilt:

$$\alpha(\lambda_1 v_1 + \dots + \lambda_k v_k) = \lambda_1 \alpha(v_1) + \dots + \lambda_k \alpha(v_k).$$

VIII.1.3. LEMMA. *Die Komposition zweier affiner Abbildungen ist affin. Die Umkehrabbildung einer bijektiven affinen Abbildung ist affin. Insbesondere bildet*

$$\text{Aff}(V) := \{\alpha: V \rightarrow V \mid \alpha \text{ affin und invertierbar}\}$$

eine Gruppe.

BEWEIS. Seien $\alpha_1: V_1 \rightarrow V_2$ und $\alpha_2: V_2 \rightarrow V_3$ zwei affine Abbildungen, $v_i \in V_1$ und $\lambda_i \in \mathbb{K}$ mit $\sum_{i=1}^k \lambda_i = 1$. Dann gilt

$$\begin{aligned} (\alpha_2 \circ \alpha_1)(\sum_{i=1}^k \lambda_i v_i) &= \alpha_2(\alpha_1(\sum_{i=1}^k \lambda_i v_i)) = \alpha_2(\sum_{i=1}^k \lambda_i \alpha_1(v_i)) \\ &= \sum_{i=1}^k \lambda_i \alpha_2(\alpha_1(v_i)) = \sum_{i=1}^k \lambda_i (\alpha_2 \circ \alpha_1)(v_i), \end{aligned}$$

also ist auch die Komposition, $\alpha_2 \circ \alpha_1: V_1 \rightarrow V_3$ affin.

Sei nun $\alpha: V \rightarrow W$ eine bijektive affine Abbildung, $w_i \in W$ und $\lambda_i \in \mathbb{K}$, sodass $\sum_{i=1}^k \lambda_i = 1$. Da α affin ist gilt:

$$\alpha(\sum_{i=1}^k \lambda_i \alpha^{-1}(w_i)) = \sum_{i=1}^k \lambda_i \alpha(\alpha^{-1}(w_i)) = \sum_{i=1}^k \lambda_i w_i.$$

Wenden wir darauf α^{-1} an, erhalten wir

$$\sum_{i=1}^k \lambda_i \alpha^{-1}(w_i) = \alpha^{-1}(\sum_{i=1}^k \lambda_i w_i),$$

also ist α^{-1} affin. □

VIII.1.4. BEISPIEL. Sei $\varphi: V \rightarrow W$ linear und $w \in W$. Dann ist

$$\alpha: V \rightarrow W, \quad \alpha(v) := \varphi(v) + w,$$

eine affine Abbildung. Sind nämlich $v_i \in V$ und $\lambda_i \in \mathbb{K}$ mit $\sum_{i=1}^k \lambda_i = 1$, so gilt

$$\begin{aligned} \alpha(\sum_{i=1}^k \lambda_i v_i) &= \varphi(\sum_{i=1}^k \lambda_i v_i) + w = \sum_{i=1}^k \lambda_i \varphi(v_i) + w \\ &= \sum_{i=1}^k \lambda_i (\varphi(v_i) + w) = \sum_{i=1}^k \lambda_i \alpha(v_i). \end{aligned}$$

Wir werden gleich sehen, dass jede affine Abbildung von dieser Form ist.

VIII.1.5. LEMMA. Ist $\alpha: V \rightarrow W$ eine affine Abbildung, dann existiert eine eindeutig bestimmte lineare Abbildung $\phi_\alpha: V \rightarrow W$ und ein eindeutiges $w_\alpha \in W$, sodass

$$\alpha(v) = \phi_\alpha(v) + w_\alpha,$$

für alle $v \in V$. Dabei ist $w_\alpha = \alpha(0)$ und $\phi_\alpha(v) = \alpha(v) - \alpha(0)$. Weiters gilt

$$\phi_{\alpha_2 \circ \alpha_1} = \phi_{\alpha_2} \phi_{\alpha_1} \quad \text{und} \quad w_{\alpha_2 \circ \alpha_1} = \phi_{\alpha_2}(w_{\alpha_1}) + w_{\alpha_2}, \quad (\text{VIII.1})$$

für je zwei affine Abbildungen $\alpha_1: V_1 \rightarrow V_2$ und $\alpha_2: V_2 \rightarrow V_3$. Schließlich ist eine affine Abbildung α genau dann invertierbar, wenn ihr linearer Teil ϕ_α invertierbar ist, und in diesem Fall gilt:

$$\phi_{\alpha^{-1}} = \phi_\alpha^{-1} \quad \text{und} \quad w_{\alpha^{-1}} = -\phi_\alpha^{-1}(w_\alpha). \quad (\text{VIII.2})$$

BEWEIS. Die Eindeutigkeitsaussage ist trivial. Es ist nur zu zeigen, dass die Abbildung $\phi_\alpha: V \rightarrow W$, $\phi_\alpha(v) := \alpha(v) - \alpha(0)$, linear ist. Sind $v_1, v_2 \in V$ so gilt

$$\begin{aligned} \phi_\alpha(v_1 + v_2) &= \alpha(v_1 + v_2) - \alpha(0) = \alpha(1 \cdot v_1 + 1 \cdot v_2 - 1 \cdot 0) - \alpha(0) \\ &= 1 \cdot \alpha(v_1) + 1 \cdot \alpha(v_2) - 1 \cdot \alpha(0) - \alpha(0) = \phi_\alpha(v_1) + \phi_\alpha(v_2), \end{aligned}$$

denn $1 \cdot v_1 + 1 \cdot v_2 - 1 \cdot 0$ ist eine Affinkombination von v_1, v_2 und 0 , da ja $1+1-1 = 1$. Für $v \in V$ und $\lambda \in \mathbb{K}$ erhalten wir analog

$$\begin{aligned} \phi_\alpha(\lambda v) &= \alpha(\lambda v) - \alpha(0) = \alpha(\lambda v + (1 - \lambda)0) - \alpha(0) \\ &= \lambda \alpha(v) + (1 - \lambda)\alpha(0) - \alpha(0) = \lambda(\alpha(v) - \alpha(0)) = \lambda \phi_\alpha(v), \end{aligned}$$

denn $\lambda v + (1 - \lambda)0$ ist eine Affinkombination von v und 0 , da ja $\lambda + (1 - \lambda) = 1$.

Sind $\alpha_1: V_1 \rightarrow V_2$ und $\alpha_2: V_2 \rightarrow V_3$ zwei affine Abbildungen so folgt

$$\begin{aligned} (\alpha_2 \circ \alpha_1)(v) &= \phi_{\alpha_2}(\alpha_1(v)) + w_{\alpha_2} \\ &= \phi_{\alpha_2}(\phi_{\alpha_1}(v) + w_{\alpha_1}) + w_{\alpha_2} = \phi_{\alpha_2}(\phi_{\alpha_1}(v)) + \phi_{\alpha_2}(w_{\alpha_1}) + w_{\alpha_2}. \end{aligned}$$

Dies zeigt (VIII.1). Daraus folgt nun auch, dass der lineare Teil einer invertierbaren affinen Abbildung invertierbar ist und, dass in diesem Fall (VIII.2) gilt, denn $\phi_{\text{id}} = \text{id}$ und $w_{\text{id}} = 0$. Ist umgekehrt α affin und ϕ_α invertierbar, dann ist $v \mapsto \phi_\alpha^{-1}(v) - \phi_\alpha^{-1}(w_\alpha)$ die Umkehrabbildung von α , also α invertierbar. \square

Nach dem vorangehenden Lemma ist

$$\text{Aff}(V) \rightarrow \text{GL}(V), \quad \alpha \mapsto \phi_\alpha, \quad (\text{VIII.3})$$

ein surjektiver Gruppenhomomorphismus. Für jedes $w \in V$ bezeichne $\tau_w: V \rightarrow V$, $\tau_w(v) := v + w$, die Translation um w . Beachte, dass dies einen injektiven Gruppenhomomorphismus

$$V \rightarrow \text{Aff}(V), \quad w \mapsto \tau_w, \quad (\text{VIII.4})$$

liefert, denn offensichtlich gilt $\tau_{w_1+w_2} = \tau_{w_1} \circ \tau_{w_2}$, und $w = \tau_w(0)$. Schließlich stimmt der Kern von (VIII.3) mit dem Bild von (VIII.4) überein, d.h.

$$\{\alpha \in \text{Aff}(V) : \phi_\alpha = \text{id}_V\} = \{\tau_v : v \in V\}.$$

Aus diesem Grund wird $\text{Aff}(V)$ als eine Gruppenerweiterung von $\text{GL}(V)$ bezeichnet. Wir erhalten eine Bijektion

$$\text{Aff}(V) \cong V \times \text{GL}(V), \quad \alpha \leftrightarrow (w_\alpha, \phi_\alpha),$$

und dies wird ein Gruppenisomorphismus, wenn wir $V \times \text{GL}(V)$ mit der Multiplikation

$$(w_2, \phi_2) \cdot (w_1, \phi_1) := (w_2 + \phi_2(w_1), \phi_2\phi_1)$$

ausstatten, siehe (VIII.1).

VIII.1.6. DEFINITION (Affine Teilräume). Eine Teilmenge A eines Vektorraums V wird *affiner Teilraum* genannt, wenn sie invariant unter Affinkombinationen ist, d.h. wenn für alle $a_i \in A$ und $\lambda_i \in \mathbb{K}$ mit $\sum_{i=1}^k \lambda_i = 1$ auch $\sum_{i=1}^k \lambda_i a_i$ in A liegt.

VIII.1.7. BEISPIEL. Die leere Menge ist ein affiner Teilraum von V .

VIII.1.8. BEISPIEL. Jeder lineare Teilraum von V ist ein affiner Teilraum.

VIII.1.9. BEISPIEL. Sind A_i affine Teilräume von V , dann ist $\bigcap_i A_i$ ein affiner Teilraum von V .

VIII.1.10. BEISPIEL. Ist $\alpha: V \rightarrow W$ eine affine Abbildung und A ein affiner Teilraum von W , dann ist das Urbild, $\alpha^{-1}(A) = \{v \in V : \alpha(v) \in A\}$, ein affiner Teilraum von V . Sind nämlich $v_i \in \alpha^{-1}(A)$ und $\lambda_i \in \mathbb{K}$ mit $\sum_{i=1}^k \lambda_i = 1$, dann folgt $\alpha(\sum_{i=1}^k \lambda_i v_i) = \sum_{i=1}^k \lambda_i \alpha(v_i) \in A$, also $\sum_{i=1}^k \lambda_i v_i \in \alpha^{-1}(A)$.

VIII.1.11. BEISPIEL. Ist $\alpha: V \rightarrow W$ eine affine Abbildung und A ein affiner Teilraum von V , dann ist das Bild, $\alpha(A)$, ein affiner Teilraum von W . Sind nämlich $w_i \in \alpha(A)$, dann existieren $a_i \in A$ mit $\alpha(a_i) = w_i$, und für $\lambda_i \in \mathbb{K}$ mit $\sum_{i=1}^k \lambda_i = 1$, erhalten wir $\sum_{i=1}^k \lambda_i w_i = \sum_{i=1}^k \lambda_i \alpha(a_i) = \alpha(\sum_{i=1}^k \lambda_i a_i) \in \alpha(A)$.

VIII.1.12. BEISPIEL. Ist W ein linearer Teilraum von V und $v \in V$, dann ist $A = v + W$ ein affiner Teilraum von V . Dies folgt aus den vorangehenden Beispielen, denn W ist affiner Teilraum und die Translation τ_v ist eine affine Abbildung. Wir werden gleich sehen, dass jeder nicht-leere affine Teilraum von dieser Gestalt ist.

VIII.1.13. LEMMA. *Ist A ein nicht-leerer affiner Teilraum von V , dann bildet*

$$V_A := \{a_2 - a_1 \mid a_1, a_2 \in A\}$$

einen linearen Teilraum von V und es gilt $A = a + V_A$, für jedes $a \in A$.

BEWEIS. Zunächst ist V_A nicht leer, denn nach Voraussetzung ist $A \neq \emptyset$. Sei $a \in A$ beliebig. Sind $v, w \in V_A$, dann existieren $a_i, b_i \in A$ mit $v = a_2 - a_1$ und $w = b_2 - b_1$, und wir erhalten

$$v + w = (1 \cdot a_2 + 1 \cdot b_2 - 1 \cdot a) - (1 \cdot a_1 + 1 \cdot b_1 - 1 \cdot a) \in V_A.$$

Ist $v \in V_A$ und $\lambda \in \mathbb{K}$, dann existieren $a_i \in A$ mit $v = a_2 - a_1$, also liegt auch

$$\lambda v = (\lambda a_2 + (1 - \lambda)a) - (\lambda a_1 + (1 - \lambda)a) \in V_A.$$

Dies zeigt, dass V_A einen linearen Teilraum von V bildet. Die Inklusion $A \subseteq a + V_A$ ist offensichtlich. Ist $v \in V_A$, dann existieren $a_i \in A$ mit $v = a_2 - a_1$ und wir erhalten $a + v = 1 \cdot a + 1 \cdot a_2 - 1 \cdot a_1 \in A$, also gilt auch die umgekehrte Inklusion, $a + V \subseteq A$. \square

VIII.1.14. DEFINITION (Dimension). Ein affiner Teilraum $A \neq \emptyset$ eines Vektorraums V wird *endlich-dimensional* genannt, wenn der lineare Teilraum V_A endlich-dimensional ist. In diesem Fall definieren wir die Dimension von A durch

$$\dim(A) := \dim(V_A).$$

1-dimensionale und 2-dimensionale affine Teilräume werden *Geraden* bzw. *Ebenen* genannt. Gilt $\dim(V) - \dim(A) = 1$, dann wird A als *affine Hyperebene* bezeichnet.

VIII.1.15. BEISPIEL. Seien A_1 und A_2 zwei affine Teilräume eines Vektorraums V , sodass $\emptyset \neq A_1 \subseteq A_2$. Dann gilt $V_{A_1} \subseteq V_{A_2}$. Ist darüber hinaus A_2 endlich-dimensional, dann ist auch A_1 endlich-dimensional und $\dim(A_1) \leq \dim(A_2)$. Ist weiters $\dim(A_1) = \dim(A_2)$, dann ist schon $A_1 = A_2$.

VIII.1.16. BEISPIEL. Ist A ein nicht-leerer affiner Teilraum eines endlich-dimensionalen Vektorraums V , und $k = \dim(V) - \dim(A)$, dann existieren affine Abbildungen $\alpha_1, \dots, \alpha_k: V \rightarrow \mathbb{K}$, sodass

$$A = \{v \in V \mid \alpha_1(v) = 0, \dots, \alpha_k(v) = 0\}.$$

Es existieren nämlich lineare Funktionale, $\varphi_1, \dots, \varphi_k: V \rightarrow \mathbb{K}$, sodass $V_A = \bigcap_{i=1}^k \ker(\varphi_i)$. Ist nun $a \in A$ beliebig, dann haben die affinen Abbildungen $\alpha_i(v) := \varphi_i(v) - \varphi_i(a)$ die gewünschte Eigenschaft.

VIII.1.17. BEISPIEL. Ist $\alpha \in \text{Aff}(V)$, dann gilt $\dim(\alpha(A)) = \dim(A)$ für jeden endlich-dimensionalen affinen Teilraum A , denn der lineare Teil von α liefert einen linearen Isomorphismus, $\phi_\alpha: V_A \rightarrow V_{\alpha(A)}$.

VIII.1.18. BEISPIEL. Sind A und A' zwei endlich-dimensionale affine Teilräume von V mit $\dim(A) = \dim(A')$, dann existiert $\alpha \in \text{Aff}(V)$, sodass $\alpha(A) = A'$. In anderen Worten, je zwei affine Teilräume gleicher Dimension sind *affin kongruent*. Nach den Ergebnissen über lineare Teilräume existiert zunächst $\phi \in \text{GL}(V)$, sodass $\phi(V_A) = V_{A'}$. Sind nun $a \in A$ und $a' \in A'$ beliebig, dann hat $\alpha(v) := \phi(v) + a' - \phi(a)$ die gewünschte Eigenschaft.

VIII.1.19. DEFINITION (Affine Hülle). Unter der *affinen Hülle* einer Teilmenge S eines Vektorraums V verstehen wir den kleinsten affinen Teilraum von V , der S enthält, d.h.

$$\langle S \rangle_{\text{aff}} := \bigcap_{S \subseteq A} A,$$

wobei der Durchschnitt über alle affinen Teilräume A von V genommen wird, die S enthalten, vgl. Beispiel VIII.1.9.

VIII.1.20. BEISPIEL. Sind $\alpha_1, \alpha_2: V \rightarrow W$ zwei affine Abbildungen und $S \subseteq V$ eine Teilmenge, sodass $\alpha_1|_S = \alpha_2|_S$, dann gilt schon $\alpha_1|_{\langle S \rangle_{\text{aff}}} = \alpha_2|_{\langle S \rangle_{\text{aff}}}$, denn $\{v \in V : \alpha_1(v) = \alpha_2(v)\}$ ist ein affiner Teilraum, der S enthält.

VIII.1.21. LEMMA. *Die affine Hülle einer Teilmenge $S \subseteq V$ lässt sich mittels Affinkombinationen wie folgt beschreiben:*

$$\langle S \rangle_{\text{aff}} = \left\{ \sum_{i=1}^k \lambda_i s_i \mid k \in \mathbb{N}, s_i \in S, \lambda_i \in \mathbb{K}, \sum_{i=1}^k \lambda_i = 1 \right\}.$$

BEWEIS. Es genügt zu zeigen, dass

$$A := \left\{ \sum_{i=1}^k \lambda_i s_i \mid k \in \mathbb{N}, s_i \in S, \lambda_i \in \mathbb{K}, \sum_{i=1}^k \lambda_i = 1 \right\}$$

einen affinen Teilraum von V bildet, denn offensichtlich gilt $\langle S \rangle_{\text{aff}} \supseteq A \supseteq S$. Seien dazu $a_1, \dots, a_l \in A$ und $\mu_1, \dots, \mu_l \in \mathbb{K}$ mit $\sum_{j=1}^l \mu_j = 1$. Es ist $\sum_{j=1}^l \mu_j a_j \in A$ zu zeigen. Da $a_j \in A$, existieren $s_j^1, \dots, s_j^{k_j} \in S$ und $\lambda_j^1, \dots, \lambda_j^{k_j} \in \mathbb{K}$ mit $\sum_{i=1}^{k_j} \lambda_j^i = 1$ und $a_j = \sum_{i=1}^{k_j} \lambda_j^i s_j^i$. Es folgt

$$\sum_{j=1}^l \mu_j a_j = \sum_{j=1}^l \mu_j \sum_{i=1}^{k_j} \lambda_j^i s_j^i = \sum_{j=1}^l \sum_{i=1}^{k_j} \mu_j \lambda_j^i s_j^i \in A,$$

denn $\sum_{j=1}^l \sum_{i=1}^{k_j} \mu_j \lambda_j^i = \sum_{j=1}^l \mu_j = 1$. \square

VIII.1.22. BEISPIEL. Ist $\alpha: V \rightarrow W$ eine affine Abbildung und $S \subseteq V$ eine Teilmenge, dann gilt

$$\alpha(\langle S \rangle_{\text{aff}}) = \langle \alpha(S) \rangle_{\text{aff}}.$$

denn einerseits ist $\alpha(\langle S \rangle_{\text{aff}})$ ein affiner Teilraum, der $\alpha(S)$ enthält, und andererseits ist $\alpha^{-1}(\langle \alpha(S) \rangle_{\text{aff}})$ ein affiner Teilraum, der S enthält.

VIII.1.23. LEMMA. *Sei $A \neq \emptyset$ ein affiner Teilraum eines Vektorraums V und $a_0, \dots, a_k \in A$. Dann sind äquivalent:*

(a) $\langle a_0, \dots, a_k \rangle_{\text{aff}} = A$

(b) Die Vektoren $a_1 - a_0, \dots, a_k - a_0$ bilden ein Erzeugendensystem von V_A .

BEWEIS. Ad (a) \Rightarrow (b): Ist $v \in V_A$, dann gilt $v + a_0 \in A = \langle a_0, \dots, a_k \rangle_{\text{aff}}$, also existieren $\lambda_0, \dots, \lambda_k \in \mathbb{K}$ mit $\sum_{i=0}^k \lambda_i = 1$, sodass $v + a_0 = \sum_{i=0}^k \lambda_i a_i$, siehe Lemma VIII.1.23. Es folgt $v = \sum_{i=0}^k \lambda_i (a_i - a_0) = \sum_{i=1}^k \lambda_i (a_i - a_0)$, also erzeugen die Vektoren $a_1 - a_0, \dots, a_k - a_0$ ganz V_A .

Ad (b) \Rightarrow (a): Es genügt $A \subseteq \langle a_0, \dots, a_k \rangle_{\text{aff}}$ zu zeigen, die umgekehrte Inklusion ist trivial. Ist $a \in A$, dann gilt $a - a_0 \in V_A = \langle a_1 - a_0, \dots, a_k - a_0 \rangle$, also existieren $\lambda_1, \dots, \lambda_k \in \mathbb{K}$ mit $a - a_0 = \sum_{i=1}^k \lambda_i (a_i - a_0)$. Wir erhalten die Darstellung $a = \sum_{i=1}^k \lambda_i a_i + (1 - \sum_{i=1}^k \lambda_i) a_0$, also $a \in \langle a_0, \dots, a_k \rangle_{\text{aff}}$. \square

VIII.1.24. BEISPIEL. Sind A_1 und A_2 zwei endlich-dimensionale affine Teilräume und $A_1 \cap A_2 \neq \emptyset$, dann sind auch $A_1 \cap A_2$ und $\langle A_1 \cup A_2 \rangle_{\text{aff}}$ endlich-dimensional und es gilt

$$\dim(A_1) + \dim(A_2) = \dim(A_1 \cap A_2) + \dim(\langle A_1 \cup A_2 \rangle_{\text{aff}}).$$

Dies folgt aus der bekannten Formel für lineare Teilräume, da $V_{A_1 \cap A_2} = V_{A_1} \cap V_{A_2}$ und $V_{\langle A_1 \cup A_2 \rangle_{\text{aff}}} = V_{A_1} + V_{A_2}$. Es ist nicht möglich dem leeren affinen Teilraum eine Dimension so zuzuordnen, dass diese Formel für beliebige endlich-dimensionale Teilräume A_1 und A_2 richtig bleibt. Diese Unannehmlichkeit löst sich in der projektiven Geometrie in Wohlgefallen auf.

VIII.1.25. DEFINITION (Affine Unabhängigkeit). Sei V ein Vektorraum. Ein System $a_0, \dots, a_k \in V$ wird *affin unabhängig* genannt wenn gilt: Sind $\lambda_i, \mu_i \in \mathbb{K}$, sodass $\sum_{i=0}^k \lambda_i = 1 = \sum_{i=0}^k \mu_i$ und $\sum_{i=0}^k \lambda_i a_i = \sum_{i=0}^k \mu_i a_i$, dann muss schon $\lambda_i = \mu_i$ gelten, für jedes $i = 0, \dots, k$.

VIII.1.26. LEMMA. Für ein System $a_0, \dots, a_k \in V$ sind äquivalent:

- (a) a_0, \dots, a_k sind affin unabhängig.
- (b) $a_1 - a_0, \dots, a_k - a_0$ sind linear unabhängig.
- (c) Für jedes $i \in \{0, \dots, k\}$ gilt $\langle a_0, \dots, a_{i-1}, a_{i+1}, \dots, a_k \rangle_{\text{aff}} \neq \langle a_0, \dots, a_k \rangle_{\text{aff}}$.

BEWEIS. Ad (a) \Rightarrow (c): Indirekt angenommen es existiert $i \in \{0, \dots, k\}$, sodass $\langle a_0, \dots, a_{i-1}, a_{i+1}, \dots, a_k \rangle_{\text{aff}} = \langle a_0, \dots, a_k \rangle_{\text{aff}}$. Dann lässt sich $a_i = 1 \cdot a_i$ auch als Affinkombination von $a_0, \dots, a_{i-1}, a_{i+1}, \dots, a_k$ schreiben, was (a) widerspricht.

Ad (c) \Rightarrow (b): Mit Lemma VIII.1.23 erhalten wir

$$\langle a_1 - a_0, \dots, a_{i-1} - a_0, a_{i+1} - a_0, \dots, a_k - a_0 \rangle \neq \langle a_1 - a_0, \dots, a_k - a_0 \rangle,$$

für jedes $i \in \{1, \dots, k\}$, also sind $a_1 - a_0, \dots, a_k - a_0$ linear unabhängig, siehe Proposition III.2.6.

Ad (b) \Rightarrow (a): Seien also $\lambda_i, \mu_i \in \mathbb{K}$, sodass $\sum_{i=0}^k \lambda_i = 1 = \sum_{i=0}^k \mu_i$ und $\sum_{i=0}^k \lambda_i a_i = \sum_{i=0}^k \mu_i a_i$. Dann gilt

$$\sum_{i=1}^k \lambda_i (a_i - a_0) = \sum_{i=1}^k \mu_i (a_i - a_0),$$

also $\lambda_i = \mu_i$, für $i = 1, \dots, k$, wegen der linearen Unabhängigkeit der Vektoren $a_1 - a_0, \dots, a_k - a_0$. Da $\sum_{i=0}^k \lambda_i = 1 = \sum_{i=0}^k \mu_i$, muss auch $\lambda_0 = \mu_0$ gelten. \square

VIII.1.27. PROPOSITION. Sei A ein k -dimensionaler affiner Teilraum eines Vektorraums V und $a_0, \dots, a_k \in A$. Dann sind äquivalent:

- (a) $\langle a_0, \dots, a_k \rangle_{\text{aff}} = A$.
- (b) a_0, \dots, a_k sind affin unabhängig.
- (c) $a_1 - a_0, \dots, a_k - a_0$ ist eine Basis von V_A .

(d) Jedes $a \in A$ lässt sich in eindeutiger Weise in der Form

$$a = \sum_{i=0}^k \lambda_i a_i$$

schreiben, wobei $\lambda_i \in \mathbb{K}$ und $\sum_{i=0}^k \lambda_i = 1$. (baryzentrische Koordinaten)

(e) Jedes $a \in A$ lässt sich in eindeutiger Weise in der Form

$$a = a_0 + \sum_{i=1}^k x_i (a_i - a_0)$$

schreiben, wobei $x_i \in \mathbb{K}$. (affine Koordinaten)

BEWEIS. Die Äquivalenz der ersten drei Aussagen folgt aus Lemma VIII.1.23 und Lemma VIII.1.26. Offensichtlich ist (d) zu (a) und (b) äquivalent. Schließlich ist auch (c) zu (e) äquivalent. \square

VIII.1.28. DEFINITION (Affine Koordinatensysteme). Sei A ein k -dimensionaler affiner Teilraum von V und $a_0, \dots, a_k \in A$. Sind die (äquivalenten) Eigenschaften in Proposition VIII.1.27 erfüllt, dann wird a_0, \dots, a_k als (geordnetes) *affines Koordinatensystem* von A bezeichnet.

Aus den Ergebnissen über lineare Teilräume folgt, dass jeder endlich-dimensionale affine Teilraum ein affines Koordinatensystem besitzt. Auch kann jede affin unabhängige Teilmenge von A zu einem affinen Koordinatensystem ergänzt werden. Schließlich enthält auch jede Teilmenge S mit $\langle S \rangle_{\text{aff}} = A$ ein affines Koordinatensystem. Wir erhalten auch sofort folgende Charakterisierung der Invertierbarkeit affiner Abbildungen:

VIII.1.29. PROPOSITION. *Eine affine Abbildung $\alpha: V \rightarrow W$ ist genau dann invertierbar, wenn sie ein (und dann jedes) affine Koordinatensystem von V auf ein affines Koordinatensystem von W abbildet.*

VIII.1.30. BEISPIEL. Die Punkte a_0, \dots, a_n bilden genau dann ein affines Koordinatensystem von V , wenn folgendes gilt: Zu jedem Vektorraum W und beliebigen $w_0, \dots, w_n \in W$ existiert eine eindeutig bestimmte affine Abbildung $\alpha: V \rightarrow W$, sodass $\alpha(a_i) = w_i$, für alle $i = 0, \dots, n$.

VIII.1.31. BEISPIEL. Sind a_0, \dots, a_n und a'_0, \dots, a'_n zwei affine Koordinatensysteme von V , dann existiert ein eindeutig bestimmtes $\alpha \in \text{Aff}(V)$, sodass $\alpha(a_i) = a'_i$, für jedes $i = 0, \dots, n$. Insbesondere sind je zwei affine Koordinatensysteme affin kongruent.

Unter gewissen Voraussetzungen lassen sich die invertierbaren affinen Abbildungen auch als jene Bijektionen charakterisieren, die Geraden auf Geraden abbilden:

VIII.1.32. SATZ (Fundamentalsatz der affinen Geometrie). *Sei V ein endlich dimensionaler reeller Vektorraum mit $\dim(V) \geq 2$. Dann ist jede Kollineation, d.h. jede Bijektion $\alpha: V \rightarrow V$, die Geraden auf Geraden abbildet, eine affine Abbildung.*

BEWEIS. Seien a_1 und $a_2 \in V$. Es genügt

$$\alpha(\lambda a_1 + (1 - \lambda)a_2) = \lambda\alpha(a_1) + (1 - \lambda)\alpha(a_2)$$

für alle $\lambda \in \mathbb{R}$ zu zeigen, mittels Induktion folgt dann sofort, dass α affin ist, siehe Aufgabe 19. O.B.d.A. dürfen wir $a_1 \neq a_2$ annehmen. Wegen der Injektivität von α ist dann auch $\alpha(a_1) \neq \alpha(a_2)$. Also muss α die Gerade durch a_1 und a_2 bijektiv auf die Gerade durch $\alpha(a_1)$ und $\alpha(a_2)$ abbilden. Es existiert daher eine eindeutige bijektive Abbildung $\chi: \mathbb{R} \rightarrow \mathbb{R}$, sodass

$$\alpha(\lambda a_1 + (1 - \lambda)a_2) = \chi(\lambda)\alpha(a_1) + (1 - \chi(\lambda))\alpha(a_2), \quad \lambda \in \mathbb{R}. \quad (\text{VIII.5})$$

Es ist also $\chi = \text{id}_{\mathbb{R}}$ zu zeigen. Zunächst gilt offensichtlich:

$$\chi(0) = 0 \quad \text{und} \quad \chi(1) = 1. \quad (\text{VIII.6})$$

Da $\dim(V) \geq 2$ existiert $a_0 \in V$, sodass die Punkte a_0, a_1, a_2 affin unabhängig sind. Beachte, dass auch $\alpha(a_0), \alpha(a_1), \alpha(a_2)$ affin unabhängig sind, denn $\alpha(a_0)$ kann wegen der Injektivität von α nicht auf der Geraden durch $\alpha(a_1)$ und $\alpha(a_2)$ liegen. Es gilt nun auch:

$$\alpha(\lambda a_0 + (1 - \lambda)a_2) = \chi(\lambda)\alpha(a_0) + (1 - \chi(\lambda))\alpha(a_2), \quad \lambda \in \mathbb{R}. \quad (\text{VIII.7})$$

Um dies einzusehen, sei o.B.d.A. $0 \neq \lambda \neq 1$ und es bezeichne g die eindeutige Gerade durch $a := \lambda a_1 + (1 - \lambda)a_2$ und $a' := \lambda a_0 + (1 - \lambda)a_2$. Nach dem Strahlensatz, siehe Aufgabe 23, schneidet g die Gerade durch a_0 und a_1 nicht. Also hat auch die Gerade durch $\alpha(a_0)$ und $\alpha(a_1)$ keinen Schnittpunkt mit $\alpha(g)$. Da $\alpha(g)$ die Gerade durch $\alpha(a_1)$ und $\alpha(a_2)$ im Punkt $\alpha(a) = \chi(\lambda)\alpha(a_1) + (1 - \chi(\lambda))\alpha(a_2)$ schneidet, folgt aus dem Strahlensatz, dass ihr Schnittpunkt mit der Geraden durch $\alpha(a_0)$ und $\alpha(a_2)$ bei $\alpha(a') = \chi(\lambda)\alpha(a_0) + (1 - \chi(\lambda))\alpha(a_2)$ liegt. Damit ist (VIII.7) gezeigt. Daraus schließen wir, dass (VIII.5) richtig bleibt, wenn wir a_1 durch einen beliebigen Punkt a'_1 auf der Geraden durch a_1 und a_2 ersetzen. Eine analoge Aussage gilt für den anderen Punkt. Für je zwei Punkte a'_1 und a'_2 auf der Geraden durch a_1 und a_2 gilt daher:

$$\alpha(\lambda a'_1 + (1 - \lambda)a'_2) = \chi(\lambda)\alpha(a'_1) + (1 - \chi(\lambda))\alpha(a'_2), \quad \lambda \in \mathbb{R}.$$

Betrachten wir $a'_1 = a_2$ und $a'_2 = a_1$ erhalten wir daraus

$$\chi(1 - \lambda) = 1 - \chi(\lambda), \quad (\text{VIII.8})$$

denn:

$$\begin{aligned} \chi(1 - \lambda)\alpha(a_1) + (1 - \chi(1 - \lambda))\alpha(a_2) &= \alpha((1 - \lambda)a_1 + (1 - (1 - \lambda))a_2) \\ &= \alpha(\lambda a_2 + (1 - \lambda)a_1) \\ &= \chi(\lambda)\alpha(a_2) + (1 - \chi(\lambda))\alpha(a_1) \end{aligned}$$

Für $a'_1 = \mu a_1 + (1 - \mu)a_2$ und $a'_2 = a_2$ erhalten wir

$$\chi(\lambda\mu) = \chi(\lambda)\chi(\mu), \quad (\text{VIII.9})$$

denn:

$$\begin{aligned} & \chi(\lambda\mu)\alpha(a_1) + (1 - \chi(\lambda\mu))\alpha(a_2) \\ &= \alpha(\lambda\mu a_1 + (1 - \lambda\mu)a_2) \\ &= \alpha(\lambda(\mu a_1 + (1 - \mu)a_2) + (1 - \lambda)a_2) \\ &= \chi(\lambda)\alpha(\mu a_1 + (1 - \mu)a_2) + (1 - \chi(\lambda))\alpha(a_2) \\ &= \chi(\lambda)(\chi(\mu)\alpha(a_1) + (1 - \chi(\mu))\alpha(a_2)) + (1 - \chi(\lambda))\alpha(a_2) \\ &= \chi(\lambda)\mu\alpha(a_1) + (1 - \chi(\lambda)\chi(\mu))\alpha(a_2) \end{aligned}$$

Insbesondere $1 = \chi(1) = \chi((-1)(-1)) = \chi(-1)\chi(-1)$, also $\chi(-1) = -1$, denn wegen der Bijektivität von χ muss $\chi(-1) \neq \chi(1) = 1$ gelten. Somit auch $\chi(-\lambda) = \chi((-1)\lambda) = \chi(-1)\chi(\lambda) = -\chi(\lambda)$, und wegen (VIII.8) daher auch $\chi(1 + \lambda) = 1 + \chi(\lambda)$. Zusammen mit (VIII.9) folgt $\chi(\lambda + \mu) = \chi(\lambda(1 + \frac{\mu}{\lambda})) = \chi(\lambda)\chi(1 + \frac{\mu}{\lambda}) = \chi(\lambda)(1 + \chi(\frac{\mu}{\lambda})) = \chi(\lambda) + \chi(\lambda)\chi(\frac{\mu}{\lambda}) = \chi(\lambda) + \chi(\mu)$, also:

$$\chi(\lambda + \mu) = \chi(\lambda) + \chi(\mu). \quad (\text{VIII.10})$$

Dies zeigt, dass χ ein Körperautomorphismus von \mathbb{R} ist.

Aus (VIII.6) und (VIII.10) folgt zunächst $\chi(n) = \chi(1 + \dots + 1) = 1 + \dots + 1 = n$ für $n \in \mathbb{N}$, und dann $\chi(n) = n$, für alle $n \in \mathbb{Z}$. Mit (VIII.9) und $0 \neq m, n \in \mathbb{Z}$ daher $n = \chi(n) = \chi(m \frac{n}{m}) = \chi(m)\chi(\frac{n}{m}) = m\chi(\frac{n}{m})$, also $\chi(q) = q$, für alle $q \in \mathbb{Q}$. Ist $\lambda \geq \mu$, dann existiert $r \in \mathbb{R}$, sodass $\lambda - \mu = r^2$, also $\chi(\lambda) - \chi(\mu) = \chi(\lambda - \mu) = \chi(r^2) = \chi(r)^2 \geq 0$, und somit auch $\chi(\lambda) \geq \chi(\mu)$. Daraus folgt nun $\chi(\lambda) = \lambda$. Wäre nämlich $\chi(\lambda) > \lambda$, dann existiert $q \in \mathbb{Q}$ mit $\chi(\lambda) > q \geq \lambda$, also auch $\chi(q) \geq \chi(\lambda)$, im Widerspruch zu $\chi(q) = q$. Analog lässt sich die Annahme $\chi(\lambda) < \lambda$ auf einen Widerspruch führen. Folglich $\chi(\lambda) = \lambda$, und der Beweis ist vollständig. \square

VIII.1.33. SATZ. *Sei V ein endlich dimensionaler Euklidischer Vektorraum und $\alpha: V \rightarrow V$ eine Abbildung. Dann sind äquivalent:*

- (a) $d(\alpha(v), \alpha(w)) = d(v, w)$, für alle $v, w \in V$.
- (b) α ist affin mit linearem Teil $\phi_\alpha \in O(V)$.
- (c) α ist affin und für ein (und dann jedes) affine Koordinatensystem a_0, \dots, a_n von V gilt $d(\alpha(a_i), \alpha(a_j)) = d(a_i, a_j)$, für alle $0 \leq i, j \leq n$.

BEWEIS. Die nicht triviale Implikation (a) \Rightarrow (b) haben wir bereits weiter oben gezeigt, siehe Satz VII.4.7. Die Implikationen (b) \Rightarrow (c) und (b) \Rightarrow (a) sind offensichtlich. Es genügt daher (c) \Rightarrow (b) zu zeigen. Betrachte die Vektoren $b_i := a_i - a_0$ und $b'_i := \alpha(a_i) - \alpha(a_0)$, $1 \leq i \leq n$. Nach Voraussetzung ist $\|b_i\| = \|b'_i\|$ und $\|b_j - b_i\| = \|b'_j - b'_i\|$, für $1 \leq i, j \leq n$. Daraus erhalten wir $\langle b_i, b_j \rangle = \langle b'_i, b'_j \rangle$, also

$$\langle \phi_\alpha(b_i), \phi_\alpha(b_j) \rangle = \langle b_i, b_j \rangle, \quad 1 \leq i, j \leq n, \quad (\text{VIII.11})$$

denn $\phi_\alpha(b_i) = b'_i$. Nach Proposition VIII.1.27 bilden die Vektoren b_1, \dots, b_n eine Basis von V . Sind $v = \sum_{i=1}^n \lambda_i v_i$ und $w = \sum_{j=1}^n \mu_j b_j$ beliebige Vektoren in V , so

$$\langle \phi_\alpha(v), \phi_\alpha(w) \rangle = \sum_{i,j=1}^n \lambda_i \mu_j \langle \phi_\alpha(b_i), \phi_\alpha(b_j) \rangle = \sum_{i,j=1}^n \lambda_i \mu_j \langle b_i, b_j \rangle = \langle v, w \rangle,$$

und daher $\phi_\alpha \in O(V)$. \square

VIII.1.34. DEFINITION (Bewegungen). Sei V ein Euklidischer Vektorraum. Eine Abbildung $\alpha: V \rightarrow V$ wird *Bewegung* genannt, wenn sie die äquivalenten Eigenschaften in Satz VIII.1.33 hat.

Offensichtlich bildet die Menge der Bewegungen eine Untergruppe von $\text{Aff}(V)$, die alle Translationen enthält.

VIII.1.35. BEISPIEL. Sind a_0, \dots, a_n und a'_0, \dots, a'_n zwei affine Koordinatensysteme eines Euklidischen Vektorraums, sodass $d(a_i, a_j) = d(a'_i, a'_j)$ für alle $0 \leq i, j \leq n$, dann existiert eine eindeutige Bewegung $\alpha: V \rightarrow V$ mit $\alpha(a_i) = a'_i$, für alle $0 \leq i \leq n$. Insbesondere sind zwei affine Koordinatensysteme a_0, \dots, a_n und a'_0, \dots, a'_n genau dann metrisch kongruent, wenn $d(a_i, a_j) = d(a'_i, a'_j)$ für alle $0 \leq i, j \leq n$.

VIII.1.36. BEISPIEL. Sind A und A' zwei affine Teilräume eines Euklidischen Vektorraums mit $\dim(A) = \dim(A')$, dann existiert eine Bewegung $\alpha: V \rightarrow V$, sodass $\alpha(A) = A'$. Ist nämlich $\phi \in O(V)$ mit $\phi(V_A) = V_{A'}$ und $a \in A$, $a' \in A'$, dann hat $\alpha(v) = \phi(v - a) + a'$ die gewünschte Eigenschaft. Zwei affine Teilräume sind also genau dann metrisch kongruent, wenn sie gleiche Dimension haben.

VIII.1.37. BEISPIEL. Sei $A \neq \emptyset$ ein affiner Teilraum eines Euklidischen Vektorraums V . Unter der Spiegelung $\sigma_A: V \rightarrow V$ an A verstehen wir die Bewegung

$$\sigma_A(v) := \sigma_{V_A}(v - a) + a,$$

wobei $\sigma_{V_A} \in O(V)$ die orthogonale Spiegelung am linearen Teilraum V_A bezeichnet und $a \in A$ beliebig ist. Beachte, dass dies nicht von der Wahl von $a \in A$ abhängt. Offensichtlich gilt $\sigma_A^2 = \text{id}_V$ und $\{v \in V : \sigma_A(v) = v\} = A$. Etwa ist die Spiegelung an einem Punkt $P \in V$ durch $\sigma_P(v) = -(v - P) + P$ gegeben.

VIII.1.38. SATZ. Sei V ein n -dimensionaler Euklidischer Vektorraum. Dann lässt sich jede Bewegung $V \rightarrow V$ als Komposition von höchstens $n + 1$ Spiegelungen an affinen Hyperebenen schreiben.

BEWEIS. In Satz VII.4.6 haben wir gesehen, dass sich jede orthogonale lineare Abbildung $V \rightarrow V$ als Komposition von höchstens n Spiegelungen an Hyperebenen durch den Koordinatenursprung schreiben lässt. In der Vorlesung haben wir

dann gezeigt, dass sich jede Translation als Komposition von zwei Spiegelungen an affinen Hyperebenen schreiben lässt.¹³ Dies zeigt aber nur, dass sich jede Bewegung als Komposition von höchstens $n + 2$ Spiegelungen schreiben lässt.

Um den Satz in voller Allgemeinheit zu beweisen sei nun $\alpha: V \rightarrow V$ eine Bewegung. O.B.d.A. dürfen wir $\alpha(0) \neq 0$ annehmen, andernfalls ist α eine lineare orthogonale Abbildung und wir kommen sogar mit n Spiegelungen aus, siehe Satz VII.4.6. Es ist dann $A := \frac{1}{2}\alpha(0) + \alpha(0)^\perp$ eine affine Hyperebene, und die Spiegelung σ_A bildet $\alpha(0)$ auf 0 ab. Somit ist $\sigma_A \circ \alpha$ eine Bewegung, die den Koordinatenursprung fixiert, d.h. eine orthogonale lineare Abbildung. Nach Satz VII.4.6 lässt sich $\sigma_A \circ \alpha$ also als Komposition von höchstens n Spiegelungen an (affinen) Hyperebenen schreiben. Folglich ist α Komposition von $n + 1$ solchen Spiegelungen. \square

VIII.2. Affine Quadriken. Sei V ein endlich dimensionaler Vektorraum über einem Körper \mathbb{K} mit Charakteristik, $\text{char}(\mathbb{K}) \neq 2$. Unter einer *quadratischen Funktion* auf V verstehen wir jede Funktion, $Q: V \rightarrow \mathbb{K}$, die sich in der Form

$$Q(v) = q(v) + l(v) + c$$

schreiben lässt, wobei $c \in \mathbb{K}$, $l \in V^*$ und q eine quadratische Form auf V bezeichnet, d.h. $q(v) = b(v, v)$ für eine (eindeutige) symmetrische Bilinearform b auf V . Beachte, dass q , l und c durch Q eindeutig bestimmt sind, denn $c = Q(0)$ und $2l(v) = Q(v) - Q(-v)$. Unter dem *Rang* einer quadratischen Funktion verstehen wir den Rang der quadratischen Form q . Im reellen Fall, $\mathbb{K} = \mathbb{R}$, wird die *Signatur* von Q als die Signatur der quadratischen Form q definiert. Unter einer *Quadrik* verstehen wir eine Teilmenge E von V , die sich in der Form

$$E = \{v \in V : Q(v) = 0\}$$

schreiben lässt, wobei Q eine quadratische Funktion auf V bezeichnet.

VIII.2.1. BEISPIEL. Für $a, b > 0$ und $x_0, y_0 \in \mathbb{R}$ bildet die durch die Gleichung

$$\left(\frac{x - x_0}{a}\right)^2 + \left(\frac{y - y_0}{b}\right)^2 = 1$$

beschriebene Ellipse mit Mittelpunkt $\begin{pmatrix} x_0 \\ y_0 \end{pmatrix}$ eine Quadrik in \mathbb{R}^2 . Analog sind auch Hyperbeln und Parabeln Quadriken in \mathbb{R}^2 .

¹³Dabei sind wir wie folgt vorgegangen: Sei $\tau_w: V \rightarrow V$, $\tau_w(v) = v + w$, eine Translation, $w \in V$. Weiters sei A_1 eine affine Hyperebene, sodass $V_{A_1} = w^\perp$, und setze $A_2 := A_1 + \frac{1}{2}w$. Dann gilt $\sigma_{V_{A_1}} = \sigma_{w^\perp} = \sigma_{V_{A_2}}$. Für $a_1 \in A_1$ und $a_2 := a_1 + \frac{1}{2}w \in A_2$ erhalten wir

$$\begin{aligned} (\sigma_{A_2} \circ \sigma_{A_1})(v) &= \sigma_{A_2}(\sigma_{w^\perp}(v - a_1) + a_1) \\ &= \sigma_{w^\perp}(\sigma_{w^\perp}(v - a_1) + a_1 - a_2) + a_2 \\ &= v - a_1 + \sigma_{w^\perp}(-\frac{1}{2}w) + a_1 + \frac{1}{2}w \\ &= v + \frac{1}{2}w + \frac{1}{2}w = \tau_w(v), \end{aligned}$$

also $\sigma_{A_2} \circ \sigma_{A_1} = \tau_w$.

VIII.2.2. BEISPIEL. Auch die Menge der Punkte $\begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2$, die der quadratischen Gleichung $x^2 - y^2 = 0$ genügen bildet eine Quadrik. Da $x^2 - y^2 = (x - y)(x + y)$ besteht sie aus zwei verschiedenen Geraden durch den Koordinatenursprung. Noch degenerierter ist die Quadrik $x^2 + y^2 + 1 = 0$, sie enthält nämlich gar keinen Punkt. Beachte, dass dieselbe Quadrik auch durch die Gleichung $x^2 + 1 = 0$ beschrieben werden kann.

VIII.2.3. LEMMA. Sei $\alpha: W \rightarrow V$ affin. Ist $Q: V \rightarrow \mathbb{K}$ eine quadratische Funktion auf V , dann ist $Q \circ \alpha$ eine quadratische Funktion auf W . Insbesondere ist für jede Quadrik E in V auch $\alpha^{-1}(E)$ eine Quadrik in W . Ist α invertierbar, dann haben $Q \circ \alpha$ und Q den selben Rang, und im reellen Fall auch die gleiche Signatur.

BEWEIS. Es sei $Q(v) = q(v) + l(v) + c$ und b die mit q assoziierte symmetrische Bilinearform, $q(w) = b(w, w)$. Weiters sei $\alpha(w) = \phi_\alpha(w) + v_\alpha$, wobei $\phi_\alpha: W \rightarrow V$ linear und $v_\alpha \in V$. Eine einfache Rechnung zeigt $Q(\alpha(w)) = q'(w) + l'(w) + c'$ wobei:

$$\begin{aligned} q'(w) &= q(\phi_\alpha(w)) \\ l'(w) &= l(\phi_\alpha(w)) + 2b(v_\alpha, \phi_\alpha(w)) \\ c' &= Q(\alpha(0)) = q(v_\alpha) + l(v_\alpha) + c \end{aligned}$$

Mit $E = \{v \in V : Q(v) = 0\}$ ist daher auch $\alpha^{-1}(E) = \{w \in W : (Q \circ \alpha)(w) = 0\}$ eine Quadrik. Alle weiteren Behauptungen sind nun offensichtlich. \square

VIII.2.4. PROPOSITION (Mittelpunkte). Ist $Q(v) = q(v) + l(v) + c$ eine quadratische Funktion auf V und $m \in V$, dann sind äquivalent:

- (a) $Q \circ \sigma_m = Q$, wobei $\sigma_m: V \rightarrow V$ die Spiegelung am Punkt m bezeichnet.
- (b) $Q \circ \tau_m = q + c'$, wo $\tau_m: V \rightarrow V$ die Translation um m bezeichnet, $c' = Q(m)$.
- (c) $Q(m + v) = q(v) + c'$, für alle $v \in V$, wobei $c' = Q(m)$.
- (d) $2b(m, v) + l(v) = 0$, für alle $v \in V$, wobei b die mit q assoziierte symmetrische Bilinearform bezeichnet, $q(v) = b(v, v)$.

Jeder Punkt m , der diese äquivalenten Eigenschaften besitzt, wird als Mittelpunkt der quadratischen Funktion Q bezeichnet. Die Menge aller Mittelpunkte, M , ist entweder leer, oder bildet einen affinen Teilraum mit $V_M = \ker(b)$ und $\dim(M) = \dim(V) - \text{rank}(q)$. Darüber hinaus ist Q auf der Menge der Mittelpunkte konstant. Ist q nicht-degeneriert, dann besitzt Q genau einen Mittelpunkt.

BEWEIS. Da $\sigma_m \circ \tau_m = \tau_m \circ \sigma_0$ ist (a) zu $Q \circ \tau_m \circ \sigma_0 = Q \circ \tau_m$ äquivalent. Dies bedeutet gerade, dass der lineare Teil der quadratischen Funktion $Q \circ \tau_m$ verschwindet, also (b). Die Äquivalenz (b) \Leftrightarrow (c) ist offensichtlich. Eine einfache Rechnung zeigt (c) \Leftrightarrow (d). Die restlichen Behauptungen folgen nun aus bekannten Eigenschaften linearer Gleichungssysteme. \square

Ist $V = \mathbb{K}^n$, so kann jede quadratische Funktion $Q: \mathbb{K}^n \rightarrow \mathbb{K}$ in der Form

$$Q(x) = x^t A x + b^t x + c, \quad x \in \mathbb{K}^n,$$

geschrieben werden, wobei $A^t = A \in M_{n \times n}(\mathbb{K})$ eine symmetrische Matrix bezeichnet, $b \in \mathbb{K}^n$ und $c \in \mathbb{K}$. Die Gleichung für Mittelpunkte $m \in \mathbb{K}^n$ von Q lautet dann $2m^t A + b^t = 0$, oder äquivalent dazu:

$$2Am + b = 0.$$

Ist A nicht degeneriert, dann ist der eindeutige Mittelpunkt daher durch

$$m = -\frac{1}{2}A^{-1}b$$

gegeben. Ist m ein Mittelpunkt, dann gilt

$$Q(m + x) = x^t Ax + c',$$

wobei $c' = Q(m)$.

VIII.2.5. BEISPIEL. Die quadratische Funktion $Q: \mathbb{R}^2 \rightarrow \mathbb{R}$

$$Q\left(\begin{pmatrix} x \\ y \end{pmatrix}\right) = x^2 + 2y^2 - 2x - 12y + 18 = \begin{pmatrix} x \\ y \end{pmatrix}^t \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} -2 \\ -12 \end{pmatrix}^t \begin{pmatrix} x \\ y \end{pmatrix} + 18,$$

besitzt genau einen Mittelpunkt, nämlich

$$m = -\frac{1}{2} \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}^{-1} \begin{pmatrix} -2 \\ -12 \end{pmatrix} = \begin{pmatrix} 1 \\ 3 \end{pmatrix}.$$

In bei m zentrierten Koordinaten hat Q daher keinen linearen Teil,

$$Q\left(m + \begin{pmatrix} x \\ y \end{pmatrix}\right) = x^2 + 2y^2 - 1.$$

Die Menge der Punkte $\begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2$, für die $Q\left(\begin{pmatrix} x \\ y \end{pmatrix}\right) = 0$ gilt, bildet daher eine Ellipse.

VIII.2.6. BEISPIEL. Die quadratische Funktion $Q: \mathbb{R}^2 \rightarrow \mathbb{R}$

$$Q\left(\begin{pmatrix} x \\ y \end{pmatrix}\right) = y^2 - x = \begin{pmatrix} x \\ y \end{pmatrix}^t \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} -1 \\ 0 \end{pmatrix}^t \begin{pmatrix} x \\ y \end{pmatrix} + 0,$$

besitzt keinen Mittelpunkt, denn das Gleichungssystem $2\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}m + \begin{pmatrix} -1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$ besitzt keine Lösung. Die Menge der Punkte $\begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2$, für die $Q\left(\begin{pmatrix} x \\ y \end{pmatrix}\right) = 0$ gilt, bildet hier eine Parabel.

VIII.2.7. BEISPIEL. Betrachte die quadratische Funktion $Q: \mathbb{R}^3 \rightarrow \mathbb{R}$

$$Q\left(\begin{pmatrix} x \\ y \\ z \end{pmatrix}\right) = x^2 + y^2 - 2x = \begin{pmatrix} x \\ y \\ z \end{pmatrix}^t \begin{pmatrix} 1 & & \\ & 1 & \\ & & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} + \begin{pmatrix} -2 \\ 0 \\ 0 \end{pmatrix}^t \begin{pmatrix} x \\ y \\ z \end{pmatrix}.$$

In diesem Fall hat die Gleichung $2Am + b$ einen 1-dimensionalen Lösungsraum, die Menge der Mittelpunkte ist $M = \left\{ \begin{pmatrix} 1 \\ 0 \\ t \end{pmatrix} : t \in \mathbb{R} \right\}$. Für jedes $m \in M$ gilt:

$$Q\left(m + \begin{pmatrix} x \\ y \\ z \end{pmatrix}\right) = x^2 + y^2 - 1.$$

Die Menge der Punkte $\begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3$, für die $Q\left(\begin{pmatrix} x \\ y \\ z \end{pmatrix}\right) = 0$ gilt, bildet hier einen kreisförmigen Zylinder mit Achse M .

VIII.2.8. SATZ (Affine Normalform quadratischer Funktionen). Sei V ein n -dimensionaler Vektorraum über einem Körper \mathbb{K} mit $\text{char}(\mathbb{K}) \neq 2$, und sei $Q: V \rightarrow \mathbb{K}$ eine quadratische Funktion. Dann existiert ein affines Koordinatensystem a_0, \dots, a_n von V , sodass die Funktion $\tilde{Q}: \mathbb{K}^n \rightarrow \mathbb{K}$,

$$\tilde{Q}(x) := Q(a_0 + \sum_{i=1}^n x_i(a_i - a_0))$$

folgende Gestalt hat:

(a) $\tilde{Q}(x) = \sum_{i=1}^r \lambda_i x_i^2 + c$, falls Q einen Mittelpunkt besitzt, $c = Q(a_0)$.

(b) $\tilde{Q}(x) = \sum_{i=1}^r \lambda_i x_i^2 + x_{r+1}$, falls Q keinen Mittelpunkt besitzt.

Dabei sind $0 \neq \lambda_i \in \mathbb{K}$, $r = \text{rank}(Q)$ und $c \in \mathbb{K}$. In anderen Worten, $\alpha: \mathbb{K}^n \rightarrow V$, $\alpha(x) := a_0 + \sum_{i=1}^n x_i(a_i - a_0)$, ist ein affiner Isomorphismus und $Q \circ \alpha$ stimmt mit einer der beiden Normalformen oben überein.

Besitzt jedes Element von \mathbb{K} eine Quadratwurzel (etwa weil \mathbb{K} algebraisch abgeschlossen ist), dann kann darüber hinaus $\lambda_i = 1$ erreicht werden. Im reellen Fall dürfen wir $\lambda_1 = \dots = \lambda_p = 1$ und $\lambda_{p+1} = \dots = \lambda_{p+q} = -1$ annehmen, wobei (p, q) die Signatur von Q bezeichnet, $p + q = r$.

BEWEIS. Wir betrachten zunächst den Fall, wo Q einen Mittelpunkt besitzt. Ist a_0 ein Mittelpunkt, dann gilt $Q(a_0 + v) = q(v) + c$, siehe Proposition VIII.2.4. Nach Satz VII.1.25 existiert eine Basis b_1, \dots, b_n von V , sodass

$$q\left(\sum_{i=1}^n x_i b_i\right) = \lambda_1 x_1^2 + \dots + \lambda_r x_r^2, \quad (\text{VIII.12})$$

wobei $r = \text{rank}(q)$ und $0 \neq \lambda_i \in \mathbb{K}$. Besitzt jedes Element in \mathbb{K} eine Quadratwurzel, dann dürfen wir $\lambda_i = 1$ annehmen, siehe Satz VII.1.25. Im reellen Fall können wir immerhin $\lambda_i = \pm 1$ erreichen, siehe Satz VII.1.35. Setzen wir $a_i := a_0 + b_i$, $i = 1, \dots, n$, dann bildet a_0, a_1, \dots, a_n das gesuchte affine Koordinatensystem, siehe Proposition VIII.1.27.

Sei nun $Q(v) = q(v) + l(v) + c$ eine quadratische Funktion, die keinen Mittelpunkt besitzt. Wie zuvor sei b_1, \dots, b_n eine Basis von V , sodass (VIII.12) und $\lambda_i \neq 0$ gilt. Da Q keinen Mittelpunkt besitzt liegt das Funktional $l \in V^*$ nicht im Bild von $b: V \rightarrow V^*$, wobei b die mit q assoziierte symmetrische Bilinearform bezeichnet, $q(v) = b(v, v)$, vgl. Proposition VIII.2.4(d). Da das Bild von b durch die Elemente b_1^*, \dots, b_r^* der zu b_1, \dots, b_n dualen Basis aufgespannt wird, existiert daher $j > r$, sodass $l(b_j) \neq 0$. O.B.d.A. dürfen wir $\mu := l(b_{r+1}) \neq 0$ und $l(b_{r+2}) = \dots = l(b_n) = 0$ annehmen. Setzen wir

$$\tilde{a}_0 := - \sum_{j=1}^r \frac{l(b_j)}{2\lambda_j} b_j,$$

dann gilt $2q(\tilde{a}_0, b_i) + l(b_i) = 0$, für alle $1 \leq i \leq r$, und daher

$$\begin{aligned} Q(\tilde{a}_0 + \sum_{i=1}^n x_i b_i) &= q\left(\sum_{i=1}^n x_i b_i\right) + 2b(\tilde{a}_0, \sum_{i=1}^n x_i b_i) + l\left(\sum_{i=1}^n x_i b_i\right) + \tilde{c} \\ &= \lambda_1 x_1^2 + \dots + \lambda_r x_r^2 + \mu x_{r+1} + \tilde{c}, \end{aligned}$$

wobei $\tilde{c} = Q(\tilde{a}_0)$. Setzen wir $a_0 := \tilde{a}_0 - \frac{\tilde{c}}{\mu}b_{r+1}$, dann ist

$$Q(a_0 + \sum_{i=1}^n x_i b_i) = \lambda_1 x_1^2 + \cdots + \lambda_r x_r^2 + \mu x_{r+1}.$$

Ersetzen wir b_{r+1} durch $\frac{1}{\mu}b_{r+1}$ dann können wir auch $\mu = 1$ erreichen. Sind nun $a_i := a_0 + b_i$, $i = 1, \dots, n$, dann bildet a_0, \dots, a_n ein affines Koordinatensystem mit den gewünschten Eigenschaften. \square

VIII.2.9. BEMERKUNG. Ist $Q: V \rightarrow \mathbb{K}$ eine quadratische Funktion und $0 \neq \lambda \in \mathbb{K}$, dann ist auch $Q' := \lambda Q: V \rightarrow \mathbb{K}$ eine quadratische Funktion, und beschreibt die selbe Quadrik, $\{v \in V : Q(v) = 0\} = \{v \in V : Q'(v) = 0\}$. Die Umkehrung gilt i.A. jedoch nicht, etwa beschreiben die Gleichungen $x^2 + 1 = 0$ und $x^2 + y^2 + 1 = 0$ die selbe (leere) Quadrik in \mathbb{R}^2 . Aber auch die Gleichungen $x^2 = 0$ und $x = 0$ beschreiben beide eine Gerade in \mathbb{R}^2 oder \mathbb{C}^2 .

VIII.2.10. KOROLLAR (Affine Normalform komplexer Quadriken). Sei V ein n -dimensionaler komplexer Vektorraum. Dann ist jede Quadrik in V zu einer der folgenden Quadriken in \mathbb{C}^n affin kongruent:

- (a) Typ $A_r: \{x_1^2 + \cdots + x_r^2 = 0\}$, wobei $0 \leq r \leq n$.
- (b) Typ $B_r: \{x_1^2 + \cdots + x_r^2 = 1\}$, wobei $0 \leq r \leq n$.
- (c) Typ $C_r: \{x_1^2 + \cdots + x_r^2 = x_{r+1}\}$, wobei $0 \leq r < n$.

D.h. es existiert ein affiner Isomorphismus, $V \cong \mathbb{C}^n$, der die Quadrik in V auf eine dieser Normalformen in \mathbb{C}^n abbildet.

BEWEIS. Sei $E = \{v \in V : Q(v) = 0\}$ eine Quadrik in V , wobei $Q: V \rightarrow \mathbb{C}$ eine quadratische Funktion bezeichnet. Nach Satz VIII.2.8 existiert ein affiner Isomorphismus $\alpha: \mathbb{C}^n \xrightarrow{\cong} V$, sodass

$$\alpha^{-1}(E) = \{x \in \mathbb{C}^n : x_1^2 + \cdots + x_r^2 + c = 0\}$$

oder

$$\alpha^{-1}(E) = \{x \in \mathbb{C}^n : x_1^2 + \cdots + x_r^2 + x_{r+1} = 0\},$$

wobei $r = \text{rank}(Q)$. Gilt im ersten Fall $c = 0$, dann ist $\alpha^{-1}(E)$ vom Typ A_r . Gilt im ersten Fall $c \neq 0$, so ist $\tilde{\alpha}: \mathbb{C}^n \xrightarrow{\cong} V$, $\tilde{\alpha}(x) := \alpha(\sqrt{-c}x)$, ein affiner Isomorphismus und nach Bemerkung VIII.2.9

$$\tilde{\alpha}^{-1}(E) = \{x \in \mathbb{C}^n : x_1^2 + \cdots + x_r^2 - 1 = 0\},$$

d.h. $\tilde{\alpha}^{-1}(E)$ ist vom Typ B_r . Im zweiten Fall ist $\tilde{\alpha}: \mathbb{C}^n \rightarrow \mathbb{C}^n$,

$$\tilde{\alpha}(x) := \alpha(x_1, \dots, x_r, -x_{r+1}, x_{r+2}, \dots, x_n),$$

einen affinen Isomorphismus, sodass

$$\tilde{\alpha}^{-1}(E) = \{x \in \mathbb{C}^n : x_1^2 + \cdots + x_r^2 - x_{r+1} = 0\},$$

d.h. $\tilde{\alpha}^{-1}(E)$ ist vom Typ C_r . \square

VIII.2.11. BEISPIEL. Jede Quadrik in \mathbb{C}^2 ist zu einer der folgenden Quadriken affin kongruent:

$A_0: \{0 = 0\}$	(die gesamte Ebene, \mathbb{C}^2)
$A_1: \{x^2 = 0\}$	(komplexe Gerade)
$A_2: \{x^2 + y^2 = 0\}$	(zwei sich schneidende komplexe Geraden)
$B_0: \{0 = 1\}$	(leer)
$B_1: \{x^2 = 1\}$	(zwei parallele komplexe Geraden)
$B_2: \{x^2 + y^2 = 1\}$	(komplexe Version des Kreises)
$C_0: \{0 = x\}$	(komplexe Gerade)
$C_1: \{x^2 = y\}$	(komplexe Version der Parabel)

VIII.2.12. BEISPIEL. Wir wollen den Typ der komplexen Quadrik

$$E := \left\{ \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{C}^2 \mid x^2 + 5y^2 + 4xy + 2x + 10y + 6 = 0 \right\}$$

bestimmen. Durch Ergänzen auf vollständige Quadrate erhalten wir

$$\begin{aligned} x^2 + 5y^2 + 4xy + 2x + 10y + 6 &= (x + 2y)^2 + y^2 + 2x + 10y + 6 \\ &= (x + 2y + 1)^2 + y^2 + 6y + 5 \\ &= (x + 2y + 1)^2 + (y + 3)^2 - 4 \\ &= 4 \left\{ \left(\frac{x + 2y + 1}{2} \right)^2 + \left(\frac{y + 3}{2} \right)^2 - 1 \right\} \end{aligned}$$

Es ist daher $\alpha: \mathbb{C}^2 \xrightarrow{\cong} \mathbb{C}^2$,

$$\alpha \begin{pmatrix} x \\ y \end{pmatrix} := \begin{pmatrix} (x + 2y + 1)/2 \\ (y + 3)/2 \end{pmatrix} = \frac{1}{2} \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \frac{1}{2} \begin{pmatrix} 1 \\ 3 \end{pmatrix},$$

ein affiner Isomorphismus und $\alpha(E) = \{x^2 + y^2 - 1 = 0\}$, d.h. E ist vom Typ B_2 , also affin kongruent zu einem komplexen Kreis.

VIII.2.13. KOROLLAR (Affine Normalform reeller Quadriken). Sei V ein n -dimensionaler reeller Vektorraum. Dann ist jede Quadrik in V zu einer der folgenden Quadriken in \mathbb{R}^n affin kongruent:

(a) Typ A_{pq} : $x_1^2 + \cdots + x_p^2 - x_{p+1}^2 - \cdots - x_{p+q}^2 = 0$, wobei $0 \leq q \leq p$ und $p+q \leq n$.

(b) Typ B_{pq} : $x_1^2 + \cdots + x_p^2 - x_{p+1}^2 - \cdots - x_{p+q}^2 = 1$, wobei $0 \leq p, q$ und $p+q \leq n$.

(c) Typ C_{pq} : $x_1^2 + \cdots + x_p^2 - x_{p+1}^2 - \cdots - x_{p+q}^2 = x_{p+q+1}$, wobei $0 \leq q \leq p$, $p+q < n$.

D.h. es existiert ein affiner Isomorphismus, $\mathbb{R}^n \cong V$, der die Quadrik in V auf eine dieser Normalformen in \mathbb{R}^n abbildet.

BEWEIS. Sei $E = \{v \in V : Q(v) = 0\}$ eine Quadrik in V , wobei $Q: V \rightarrow \mathbb{R}$ eine quadratische Funktion bezeichnet. Nach Satz VIII.2.8 existiert ein affiner Isomorphismus $\alpha: \mathbb{R}^n \xrightarrow{\cong} V$, sodass

$$\alpha^{-1}(E) = \{x \in \mathbb{R}^n : x_1^2 + \cdots + x_p^2 - x_{p+1}^2 - \cdots - x_{p+q}^2 + c = 0\}$$

oder

$$\alpha^{-1}(E) = \{x \in \mathbb{R}^n : x_1^2 + \cdots + x_p^2 - x_{p+1}^2 - \cdots - x_{p+q}^2 + x_{p+q+1} = 0\},$$

wobei (p, q) die Signatur von Q bezeichnet. Gilt im ersten Fall $c = 0$ und $p \geq q$, dann ist $\alpha^{-1}(E)$ vom Typ A_{pq} . Ist $p < q$, dann bildet $\tilde{\alpha}: \mathbb{R}^n \rightarrow V$,

$$\tilde{\alpha}(x) = \alpha(x_{q+1}, \dots, x_{q+p}, x_1, \dots, x_q, x_{p+q+1}, \dots, x_n),$$

einen affinen Isomorphismus, sodass

$$\tilde{\alpha}^{-1}(E) = \{x \in \mathbb{R}^n : x_1^2 + \dots + x_q^2 - x_{q+1}^2 - \dots - x_{q+p}^2 = 0\},$$

d.h. $\tilde{\alpha}^{-1}(E)$ ist vom Typ A_{qp} . Gilt im ersten Fall $c < 0$, dann ist $\tilde{\alpha}: \mathbb{R}^n \rightarrow V$, $\tilde{\alpha}(x) := \alpha(\sqrt{-c}x)$, ein affiner Isomorphismus, sodass

$$\tilde{\alpha}^{-1}(E) = \{x \in \mathbb{R}^n : x_1^2 + \dots + x_p^2 - x_{p+1}^2 - \dots - x_{p+q}^2 - 1 = 0\},$$

d.h. $\tilde{\alpha}^{-1}(E)$ ist vom Typ B_{pq} . Ist $c > 0$, dann bildet $\tilde{\alpha}: \mathbb{R}^n \rightarrow V$,

$$\tilde{\alpha}(x) := \alpha(\sqrt{c}x_{q+1}, \dots, \sqrt{c}x_{q+p}, \sqrt{c}x_1, \dots, \sqrt{c}x_q, x_{p+q+1}, \dots, x_n),$$

einen affinen Isomorphismus, sodass

$$\tilde{\alpha}^{-1}(E) = \{x \in \mathbb{R}^n : x_1^2 + \dots + x_q^2 - x_{q+1}^2 - \dots - x_{q+p}^2 - 1 = 0\},$$

d.h. $\tilde{\alpha}^{-1}(E)$ ist vom Typ B_{qp} . Gilt im zweiten Fall $p \geq q$, so ist $\tilde{\alpha}: \mathbb{R}^n \rightarrow \mathbb{R}^n$,

$$\tilde{\alpha}(x) := \alpha(x_1, \dots, x_{p+q}, -x_{p+q+1}, x_{p+q+2}, \dots, x_n),$$

ein affiner Isomorphismus mit

$$\tilde{\alpha}^{-1}(E) = \{x \in \mathbb{R}^n : x_1^2 + \dots + x_p^2 - x_{p+1}^2 - \dots - x_{p+q}^2 - x_{q+p+1} = 0\},$$

d.h. $\tilde{\alpha}^{-1}(E)$ ist vom Typ C_{pq} . Ist $p < q$, dann bildet $\tilde{\alpha}: \mathbb{R}^n \rightarrow \mathbb{R}^n$,

$$\tilde{\alpha}(x) := \alpha(x_{p+1}, \dots, x_{p+q}, x_1, \dots, x_p, x_{p+q+1}, \dots, x_n),$$

einen affinen Isomorphismus, sodass

$$\tilde{\alpha}^{-1}(E) = \{x \in \mathbb{R}^n : x_1^2 + \dots + x_q^2 - x_{q+1}^2 - \dots - x_{q+p}^2 - x_{q+p+1} = 0\},$$

d.h. $\tilde{\alpha}^{-1}(E)$ ist vom Typ C_{qp} . □

VIII.2.14. BEISPIEL. Jede Quadrik in \mathbb{R}^2 ist zu einer der folgenden Quadriken affin kongruent:

$A_{00}: \{0 = 0\}$	(die gesamte Ebene, \mathbb{R}^2)
$A_{10}: \{x^2 = 0\}$	(eine Gerade)
$A_{20}: \{x^2 + y^2 = 0\}$	(ein Punkt)
$A_{11}: \{x^2 - y^2 = 0\}$	(zwei sich schneidende Geraden)
$B_{00}: \{0 = 1\}$	(leer)
$B_{10}: \{x^2 = 1\}$	(zwei parallele Geraden)
$B_{01}: \{-x^2 = 1\}$	(leer)
$B_{20}: \{x^2 + y^2 = 1\}$	(Kreis)
$B_{11}: \{x^2 - y^2 = 1\}$	(Hyperbel)
$B_{02}: \{-x^2 - y^2 = 1\}$	(leer)
$C_{00}: \{0 = x\}$	(eine Gerade)
$C_{10}: \{x^2 = y\}$	(Parabel)

VIII.2.15. BEISPIEL. Wir wollen den Typ der reellen Quadrik

$$E := \left\{ \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2 : x^2 + y^2 + 2xy + 2x + 11y + 10 = 0 \right\}$$

bestimmen. Durch Ergänzen auf vollständige Quadrate erhalten wir:

$$\begin{aligned} x^2 + y^2 + 2xy + 2x + 11y + 10 &= (x + y)^2 + 2x + 11y + 10 \\ &= (x + y + 1)^2 + 9y + 9 \\ &= (x + y + 1)^2 + 9(y + 1) \\ &= 9 \left\{ \left(\frac{x + y + 1}{3} \right)^2 + (y + 1) \right\} \end{aligned}$$

Es ist daher, $\alpha: \mathbb{R}^2 \xrightarrow{\cong} \mathbb{R}^2$,

$$\alpha \begin{pmatrix} x \\ y \end{pmatrix} := \begin{pmatrix} (x + y + 1)/3 \\ y + 1 \end{pmatrix} = \begin{pmatrix} 1/3 & 1/3 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} 1/3 \\ 1 \end{pmatrix},$$

ein affiner Isomorphismus, und $\alpha(E) = \left\{ \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2 : x^2 + y = 0 \right\}$, d.h. E ist eine Parabel.

VIII.2.16. BEISPIEL. Jede Quadrik in \mathbb{R}^3 ist zu einer der folgenden Quadriken affin kongruent:

$A_{00}: \{0 = 0\}$	(der gesamte Raum, \mathbb{R}^3)
$A_{10}: \{x^2 = 0\}$	(Ebene)
$A_{20}: \{x^2 + y^2 = 0\}$	(Gerade)
$A_{11}: \{x^2 - y^2 = 0\}$	(zwei sich schneidende Ebenen)
$A_{30}: \{x^2 + y^2 + z^2 = 0\}$	(ein Punkt)
$A_{21}: \{x^2 + y^2 - z^2 = 0\}$	(kreisförmiger Kegel)
$B_{00}: \{0 = 1\}$	(leer)
$B_{10}: \{x^2 = 1\}$	(zwei parallele Ebenen)
$B_{01}: \{-x^2 = 1\}$	(leer)
$B_{20}: \{x^2 + y^2 = 1\}$	(kreisförmiger Zylinder)
$B_{11}: \{x^2 - y^2 = 1\}$	(hyperbolischer Zylinder)
$B_{02}: \{-x^2 - y^2 = 1\}$	(leer)
$B_{30}: \{x^2 + y^2 + z^2 = 1\}$	(Sphäre)
$B_{21}: \{x^2 + y^2 - z^2 = 1\}$	(einschaliges Hyperboloid)
$B_{12}: \{x^2 - y^2 - z^2 = 1\}$	(zweischaliges Hyperboloid)
$B_{03}: \{-x^2 - y^2 - z^2 = 1\}$	(leer)
$C_{00}: \{0 = x\}$	(Ebene)
$C_{10}: \{x^2 = y\}$	(parabolischer Zylinder)
$C_{20}: \{x^2 + y^2 = z\}$	(Paraboloid)
$C_{11}: \{x^2 - y^2 = z\}$	(hyperbolisches Paraboloid, Sattel)

VIII.2.17. BEISPIEL. Wir wollen den Typ der Quadrik $E \subseteq \mathbb{R}^3$ bestimmen, die durch die quadratische Gleichung

$$4x^2 - 8y^2 + z^2 - 4xy - 4xz - 4yz + 4x + 4y - 6z + 5 = 0$$

gegeben ist. Durch Ergänzen auf vollständige Quadrate erhalten wir:

$$\begin{aligned}
4x^2 - 8y^2 + z^2 - 4xy - 4xz - 4yz + 4x + 4y - 6z + 5 & \\
= (2x - y - z)^2 - 9y^2 - 6yz + 4x + 4y - 6z + 5 & \\
= (2x - y - z)^2 - (3y + z)^2 + z^2 + 4x + 4y - 6z + 5 & \\
= (2x - y - z + 1)^2 - (3y + z)^2 + z^2 + 6y - 4z + 4 & \\
= (2x - y - z + 1)^2 - (3y + z - 1)^2 + z^2 - 6z + 5 & \\
= (2x - y - z + 1)^2 - (3y + z - 1)^2 + (z - 3)^2 - 4 & \\
= 4 \left\{ \left(\frac{2x - y - z + 1}{2} \right)^2 + \left(\frac{z - 3}{2} \right)^2 - \left(\frac{3y + z - 1}{2} \right)^2 - 1 \right\} &
\end{aligned}$$

Es ist daher, $\alpha: \mathbb{R}^3 \rightarrow \mathbb{R}^3$,

$$\alpha \begin{pmatrix} x \\ y \\ z \end{pmatrix} := \begin{pmatrix} (2x - y - z + 1)/2 \\ (z - 3)/2 \\ (3y + z - 1)/2 \end{pmatrix} = \frac{1}{2} \begin{pmatrix} 2 & -1 & -1 \\ 0 & 0 & 1 \\ 0 & 3 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} + \frac{1}{2} \begin{pmatrix} 1 \\ -3 \\ -1 \end{pmatrix},$$

ein affiner Isomorphismus und $\alpha(E) = \{x^2 + y^2 - z^2 - 1 = 0\}$, d.h. E ist ein einschaliges Hyperboloid.

Punkte a_0, \dots, a_n eines Euklidischen Vektorraums V werden als *affines Orthonormalsystem* bezeichnet, falls $a_1 - a_0, \dots, a_n - a_0$ eine Orthonormalbasis von V bilden. Dies ist genau dann der Fall, wenn $\mathbb{R}^n \rightarrow V$, $x \mapsto a_0 + \sum_{i=1}^n x_i(a_i - a_0)$, eine Bewegung ist, d.h. Längen bewahrt. Bewegungen bilden affine Orthonormalsysteme auf affine Orthonormalsysteme ab. Umgekehrt sind je zwei affine Orthonormalsysteme metrisch kongruent, d.h. können durch eine Bewegung ineinander übergeführt werden.

VIII.2.18. SATZ (Metrische Normalform quadratischer Funktionen). *Sei V ein n -dimensionaler Euklidischer Vektorraum und $Q: V \rightarrow \mathbb{R}$ eine quadratische Funktion. Dann existiert ein affines Orthonormalsystem a_0, \dots, a_n von V , sodass die Funktion $\tilde{Q}: \mathbb{R}^n \rightarrow \mathbb{R}$, $\tilde{Q}(x_1, \dots, x_n) := Q(a_0 + \sum_{i=1}^n x_i(a_i - a_0))$ folgende Gestalt hat:*

- (a) $\tilde{Q}(x_1, \dots, x_n) = \sum_{i=1}^p \lambda_i x_i^2 - \sum_{i=p+1}^{p+q} \lambda_i x_i^2 + c$, falls Q einen Mittelpunkt besitzt.
(b) $\tilde{Q}(x_1, \dots, x_n) = \sum_{i=1}^p \lambda_i x_i^2 - \sum_{i=p+1}^{p+q} \lambda_i x_i^2 + \mu x_{p+q+1}$, falls Q keinen MP hat.

Dabei ist (p, q) die Signatur von Q , $0 < \lambda_1 \leq \dots \leq \lambda_p$, $0 < \lambda_{p+1} \leq \dots \leq \lambda_{p+q}$, $\mu > 0$ und $c \in \mathbb{R}$. In anderen Worten, $\alpha: \mathbb{R}^n \rightarrow V$, $\alpha(x) := a_0 + \sum_{i=1}^n x_i(a_i - a_0)$, ist eine Bewegung und $Q \circ \alpha$ stimmt mit einer der beiden Normalformen überein.

BEWEIS. Wir gehen genau wie im Beweis von Satz VIII.2.8 vor, können nun aber b_1, \dots, b_n als Orthonormalbasis wählen, siehe Korollar VII.3.16. \square

VIII.2.19. KOROLLAR (Metrische Normalform reeller Quadriken). Sei V ein n -dimensionaler Euklidischer Vektorraum. Dann ist jede Quadrik in V zu einer der folgenden Quadriken in \mathbb{R}^n metrisch kongruent:

(a) Typ A_{pq} :

$$\left(\frac{x_1}{a_1}\right)^2 + \cdots + \left(\frac{x_p}{a_p}\right)^2 - \left(\frac{x_{p+1}}{a_{p+1}}\right)^2 - \cdots - \left(\frac{x_{p+q}}{a_{p+q}}\right)^2 = 0,$$

wobei $0 \leq q \leq p$, $p+q \leq n$, $1 = a_1 \geq \cdots \geq a_p > 0$, und $a_{p+1} \geq \cdots \geq a_{p+q} > 0$.

(b) Typ B_{pq} :

$$\left(\frac{x_1}{a_1}\right)^2 + \cdots + \left(\frac{x_p}{a_p}\right)^2 - \left(\frac{x_{p+1}}{a_{p+1}}\right)^2 - \cdots - \left(\frac{x_{p+q}}{a_{p+q}}\right)^2 = 1,$$

wobei $0 \leq p$, $0 \leq q$, $p+q \leq n$, $a_1 \geq \cdots \geq a_p > 0$, und $a_{p+1} \geq \cdots \geq a_{p+q} > 0$.

(c) Typ C_{pq} :

$$\left(\frac{x_1}{a_1}\right)^2 + \cdots + \left(\frac{x_p}{a_p}\right)^2 - \left(\frac{x_{p+1}}{a_{p+1}}\right)^2 - \cdots - \left(\frac{x_{p+q}}{a_{p+q}}\right)^2 = x_{p+q+1},$$

wobei $0 \leq q \leq p$, $p+q < n$, $a_1 \geq \cdots \geq a_p > 0$, und $a_{p+1} \geq \cdots \geq a_{p+q} > 0$.

D.h. es existiert eine Bewegung, $\mathbb{R}^n \cong V$, die die Quadrik in V auf eine dieser Normalformen in \mathbb{R}^n abbildet. Jede solche Bewegung wird als Hauptachsentransformation bezeichnet.

BEWEIS. Sei $E = \{v \in V : Q(v) = 0\}$ eine Quadrik in V , wobei $Q: V \rightarrow \mathbb{R}$ eine quadratische Funktion bezeichnet. Nach Satz VIII.2.18 existiert eine Bewegung, $\alpha: \mathbb{R}^n \xrightarrow{\cong} V$, sodass $\alpha^{-1}(E)$ die Gestalt

$$\left\{ x \in \mathbb{R}^n \mid \left(\frac{x_1}{a_1}\right)^2 + \cdots + \left(\frac{x_p}{a_p}\right)^2 - \left(\frac{x_{p+1}}{a_{p+1}}\right)^2 - \cdots - \left(\frac{x_{p+q}}{a_{p+q}}\right)^2 + c = 0 \right\}$$

oder

$$\left\{ x \in \mathbb{R}^n \mid \left(\frac{x_1}{a_1}\right)^2 + \cdots + \left(\frac{x_p}{a_p}\right)^2 - \left(\frac{x_{p+1}}{a_{p+1}}\right)^2 - \cdots - \left(\frac{x_{p+q}}{a_{p+q}}\right)^2 + \mu x_{p+q+1} = 0 \right\}$$

hat, wobei $a_1 \geq \cdots \geq a_p > 0$, $a_{p+1} \geq \cdots \geq a_{p+q} > 0$, $c \in \mathbb{R}$ und $\mu \neq 0$. Dabei ist $a_i = 1/\sqrt{\lambda_i}$, vgl. Satz VIII.2.18. Durch einfache Manipulationen lassen sich diese Quadriken nun auf einen der Typen A_{pq} , B_{pq} oder C_{pq} transformieren.

Wir gehen genau wie im Beweis von Korollar VIII.2.13. Gilt im ersten Fall $c = 0$ und $p \geq q$, dann können wir die definierende Gleichung mit a_1^2 multiplizieren und sehen, dass $\alpha^{-1}(E)$ vom Typ A_{pq} . Ist $p < q$, dann bildet $\tilde{\alpha}: \mathbb{R}^n \rightarrow V$,

$$\tilde{\alpha}(x) = \alpha(x_{q+1}, \dots, x_{q+p}, x_1, \dots, x_q, x_{p+q+1}, \dots, x_n), \quad (\text{VIII.13})$$

eine Bewegung, sodass

$$\tilde{\alpha}^{-1}(E) = \left\{ x \in \mathbb{R}^n \mid \left(\frac{x_1}{a_{p+1}}\right)^2 + \cdots + \left(\frac{x_q}{a_{p+q}}\right)^2 - \left(\frac{x_{q+1}}{a_1}\right)^2 - \cdots - \left(\frac{x_{q+p}}{a_p}\right)^2 = 0 \right\}.$$

Multiplizieren wir die definierende Gleichung mit a_{p+1}^2 , so sehen wir, dass $\tilde{\alpha}^{-1}(E)$ vom Typ A_{qp} ist. Gilt im ersten Fall $c < 0$, dann

$$\alpha^{-1}(E) = \left\{ x \in \mathbb{R}^n \mid \left(\frac{x_1}{\tilde{a}_1}\right)^2 + \dots + \left(\frac{x_p}{\tilde{a}_p}\right)^2 - \left(\frac{x_{p+1}}{\tilde{a}_{p+1}}\right)^2 - \dots - \left(\frac{x_{p+q}}{\tilde{a}_{p+q}}\right)^2 - 1 = 0 \right\},$$

wobei $\tilde{a}_i = \sqrt{-ca_i}$, d.h. $\alpha^{-1}(E)$ ist vom Typ B_{pq} . Gilt $c > 0$, dann ist (VIII.13) eine Bewegung, sodass

$$\tilde{\alpha}^{-1}(E) = \left\{ x \in \mathbb{R}^n \mid \left(\frac{x_1}{\tilde{a}_{p+1}}\right)^2 + \dots + \left(\frac{x_q}{\tilde{a}_{p+q}}\right)^2 - \left(\frac{x_{q+1}}{\tilde{a}_1}\right)^2 - \dots - \left(\frac{x_{q+p}}{\tilde{a}_p}\right)^2 - 1 = 0 \right\}.$$

wobei $\tilde{a}_i = \sqrt{ca_i}$, d.h. $\tilde{\alpha}^{-1}(E)$ ist vom Typ B_{qp} . Auch im zweiten Fall dürfen wir o.B.d.A. $p \geq q$ annehmen in dem wir gegebenenfalls eine Bewegung der Form (VIII.13) anwenden. Durch eine Bewegung der Form

$$\tilde{\alpha}(x) := \alpha(x_1, \dots, x_{p+q}, -x_{p+q+1}, x_{p+q+2}, \dots, x_n),$$

können wir auch $\mu < 0$ erreichen. Es gilt dann

$$\tilde{\alpha}^{-1}(E) = \left\{ x \in \mathbb{R}^n \mid \left(\frac{x_1}{\tilde{a}_1}\right)^2 + \dots + \left(\frac{x_p}{\tilde{a}_p}\right)^2 - \left(\frac{x_{p+1}}{\tilde{a}_{p+1}}\right)^2 - \dots - \left(\frac{x_{p+q}}{\tilde{a}_{p+q}}\right)^2 - x_{p+q+1} = 0 \right\}.$$

wobei $\tilde{a}_i = \sqrt{-\mu a_i}$, d.h. $\alpha^{-1}(E)$ ist vom Typ C_{pq} . \square

VIII.2.20. BEISPIEL. Jede Quadrik in \mathbb{R}^2 ist zu einer der folgenden Quadriken metrisch kongruent:

$A_{00}: \{0 = 0\}$	(die gesamte Ebene, \mathbb{R}^2)
$A_{10}: \{x^2 = 0\}$	(Gerade)
$A_{20}: \{x^2 + \left(\frac{y}{b}\right)^2 = 0\}$	(ein Punkt, $b > 0$)
$A_{11}: \{x^2 - \left(\frac{y}{b}\right)^2 = 0\}$	(zwei sich schneidende Geraden, $b > 0$)
$B_{00}: \{0 = 1\}$	(leer)
$B_{10}: \left\{\left(\frac{x}{a}\right)^2 = 1\right\}$	(zwei parallele Geraden, $a > 0$)
$B_{01}: \left\{-\left(\frac{x}{a}\right)^2 = 1\right\}$	(leer, $a > 0$)
$B_{20}: \left\{\left(\frac{x}{a}\right)^2 + \left(\frac{y}{b}\right)^2 = 1\right\}$	(Ellipse in Hauptlage, $a \geq b > 0$)
$B_{11}: \left\{\left(\frac{x}{a}\right)^2 - \left(\frac{y}{b}\right)^2 = 1\right\}$	(Hyperbel in Hauptlage, $a, b > 0$)
$B_{02}: \left\{-\left(\frac{x}{a}\right)^2 - \left(\frac{y}{b}\right)^2 = 1\right\}$	(leer, $a, b > 0$)
$C_{00}: \{0 = x\}$	(Gerade)
$C_{10}: \left\{\left(\frac{x}{a}\right)^2 = y\right\}$	(Parabel in Hauptlage, $a > 0$)

VIII.2.21. BEISPIEL. Betrachte die Quadrik

$$E = \left\{ \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2 : 13x^2 + 13y^2 - 10xy + 26\sqrt{2}x - 10\sqrt{2}y - 46 = 0 \right\}.$$

Wir wollen eine Bewegung $\mathbb{R}^2 \rightarrow \mathbb{R}^2$ bestimmen, die E auf Normalform bringt. Dazu diagonalisieren wir zunächst den quadratischen Teil der Gleichung:

$$\begin{pmatrix} 13 & -5 \\ -5 & 13 \end{pmatrix} = U \begin{pmatrix} 8 & 0 \\ 0 & 18 \end{pmatrix} U^t, \quad \text{wobei } U = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} \in O_2.$$

Setzen wir $\begin{pmatrix} \tilde{x} \\ \tilde{y} \end{pmatrix} := U^t \begin{pmatrix} x \\ y \end{pmatrix}$ dann folgt für den quadratischen Teil,

$$13x^2 + 13y^2 - 10xy = \begin{pmatrix} x \\ y \end{pmatrix}^t \begin{pmatrix} 13 & -5 \\ -5 & 13 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \tilde{x} \\ \tilde{y} \end{pmatrix}^t \begin{pmatrix} 8 & 0 \\ 0 & 18 \end{pmatrix} \begin{pmatrix} \tilde{x} \\ \tilde{y} \end{pmatrix} = 8\tilde{x}^2 + 18\tilde{y}^2,$$

und daher mit $\begin{pmatrix} x \\ y \end{pmatrix} = U \begin{pmatrix} \tilde{x} \\ \tilde{y} \end{pmatrix} = \begin{pmatrix} (\tilde{x}-\tilde{y})/\sqrt{2} \\ (\tilde{x}+\tilde{y})/\sqrt{2} \end{pmatrix}$:

$$\begin{aligned} 13x^2 + 13y^2 - 10xy + 26\sqrt{2}x - 10\sqrt{2}y - 46 &= 8\tilde{x}^2 + 18\tilde{y}^2 + 26\sqrt{2}(\tilde{x} - \tilde{y})/\sqrt{2} - 10\sqrt{2}(\tilde{x} + \tilde{y})/\sqrt{2} - 36\tilde{y} - 46 \\ &= 8\tilde{x}^2 + 18\tilde{y}^2 + 16\tilde{x} - 36\tilde{y} - 46 \\ &= 8(\tilde{x} + 1)^2 + 18(\tilde{y} - 1)^2 - 72 \\ &= 72 \left\{ \left(\frac{\tilde{x} + 1}{3} \right)^2 + \left(\frac{\tilde{y} - 1}{2} \right)^2 - 1 \right\} \end{aligned}$$

Somit ist $\alpha: \mathbb{R}^2 \rightarrow \mathbb{R}^2$,

$$\alpha \begin{pmatrix} x \\ y \end{pmatrix} := U^t \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} 1 \\ -1 \end{pmatrix} = \begin{pmatrix} \tilde{x} + 1 \\ \tilde{y} - 1 \end{pmatrix} = \begin{pmatrix} (x + y)/\sqrt{2} + 1 \\ (-x + y)/\sqrt{2} - 1 \end{pmatrix},$$

eine Bewegung, für die

$$\alpha(E) = \left\{ \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2 : \left(\frac{x}{3} \right)^2 + \left(\frac{y}{2} \right)^2 = 1 \right\}$$

gilt. Es handelt sich also um eine Ellipse mit Halbachsen 3 und 2.

VIII.2.22. BEISPIEL. Jede Quadrik in \mathbb{R}^3 ist zu einer der folgenden Quadriken metrisch kongruent:

$A_{00}: \{0 = 0\}$	(der gesamte Raum, \mathbb{R}^3)
$A_{10}: \{x^2 = 0\}$	(Ebene)
$A_{20}: \{x^2 + \left(\frac{y}{b}\right)^2 = 0\}$	(Gerade, $b > 0$)
$A_{11}: \{x^2 - \left(\frac{y}{b}\right)^2 = 0\}$	(zwei sich schneidende Ebenen, $b > 0$)
$A_{30}: \{x^2 + \left(\frac{y}{b}\right)^2 + \left(\frac{z}{c}\right)^2 = 0\}$	(ein Punkt, $b, c > 0$)
$A_{21}: \{x^2 + \left(\frac{y}{b}\right)^2 - \left(\frac{z}{c}\right)^2 = 0\}$	(Kegel, $b, c > 0$)
$B_{00}: \{0 = 1\}$	(leer)
$B_{10}: \left\{ \left(\frac{x}{a} \right)^2 = 1 \right\}$	(zwei parallele Ebenen, $a > 0$)
$B_{01}: \left\{ - \left(\frac{x}{a} \right)^2 = 1 \right\}$	(leer, $a > 0$)
$B_{20}: \left\{ \left(\frac{x}{a} \right)^2 + \left(\frac{y}{b} \right)^2 = 1 \right\}$	(elliptischer Zylinder, $a \geq b > 0$)
$B_{11}: \left\{ \left(\frac{x}{a} \right)^2 - \left(\frac{y}{b} \right)^2 = 1 \right\}$	(hyperbolischer Zylinder, $a, b > 0$)
$B_{02}: \left\{ - \left(\frac{x}{a} \right)^2 - \left(\frac{y}{b} \right)^2 = 1 \right\}$	(leer, $a, b > 0$)
$B_{30}: \left\{ \left(\frac{x}{a} \right)^2 + \left(\frac{y}{b} \right)^2 + \left(\frac{z}{c} \right)^2 = 1 \right\}$	(Ellipsoid, $a \geq b \geq c > 0$)
$B_{21}: \left\{ \left(\frac{x}{a} \right)^2 + \left(\frac{y}{b} \right)^2 - \left(\frac{z}{c} \right)^2 = 1 \right\}$	(einschal. Hyperboloid, $a \geq b > 0, c > 0$)
$B_{12}: \left\{ \left(\frac{x}{a} \right)^2 - \left(\frac{y}{b} \right)^2 - \left(\frac{z}{c} \right)^2 = 1 \right\}$	(zweischal. Hyperboloid, $a > 0, b \geq c > 0$)
$B_{03}: \left\{ - \left(\frac{x}{a} \right)^2 - \left(\frac{y}{b} \right)^2 - \left(\frac{z}{c} \right)^2 = 1 \right\}$	(leer, $a \geq b \geq c > 0$)

$$\begin{aligned}
C_{00}: \{0 = x\} & \qquad \qquad \qquad \text{(Ebene)} \\
C_{10}: \left\{ \left(\frac{x}{a}\right)^2 = y \right\} & \qquad \qquad \text{(parabolischer Zylinder, } a > 0) \\
C_{20}: \left\{ \left(\frac{x}{a}\right)^2 + \left(\frac{y}{b}\right)^2 = z \right\} & \qquad \text{(elliptisches Paraboloid, } a \geq b > 0) \\
C_{11}: \left\{ \left(\frac{x}{a}\right)^2 - \left(\frac{y}{b}\right)^2 = z \right\} & \qquad \text{(hyperbolisches Paraboloid, Sattel, } a, b > 0)
\end{aligned}$$

VIII.2.23. BEISPIEL. Betrachte die Quadrik $E \subseteq \mathbb{R}^3$, die durch die Gleichung

$$-21x^2 - 5y^2 - 5z^2 + 30xy + 30xz - 2yz + 8\sqrt{2}y - 8\sqrt{2}z - 52 = 0$$

beschrieben ist. Wir wollen eine Bewegung $\alpha: \mathbb{R}^3 \rightarrow \mathbb{R}^3$ bestimmen, die E auf Normalform bringt. Wir diagonalisieren zunächst den quadratischen Teil der Gleichung und bestimmen dazu eine orthogonale Matrix U , sodass

$$\begin{pmatrix} -21 & 15 & 15 \\ 15 & -5 & -1 \\ 15 & -1 & -5 \end{pmatrix} = U \begin{pmatrix} 9 & & \\ & -4 & \\ & & -36 \end{pmatrix} U^t,$$

mit einer orthogonalen Matrix,

$$U = \begin{pmatrix} 1/\sqrt{3} & 0 & -2/\sqrt{6} \\ 1/\sqrt{3} & -1/\sqrt{2} & 1/\sqrt{6} \\ 1/\sqrt{3} & 1/\sqrt{2} & 1/\sqrt{6} \end{pmatrix} \in O_3.$$

Setzen wir $\begin{pmatrix} \tilde{x} \\ \tilde{y} \\ \tilde{z} \end{pmatrix} := U^t \begin{pmatrix} x \\ y \\ z \end{pmatrix}$ dann erhalten wir für den quadratischen Teil:

$$\begin{aligned}
-21x^2 - 5y^2 - 5z^2 + 30xy + 30xz - 2yz &= \begin{pmatrix} x \\ y \\ z \end{pmatrix}^t \begin{pmatrix} -21 & 15 & 15 \\ 15 & -5 & -1 \\ 15 & -1 & -5 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} \\
&= \begin{pmatrix} \tilde{x} \\ \tilde{y} \\ \tilde{z} \end{pmatrix}^t \begin{pmatrix} 9 & & \\ & -4 & \\ & & -36 \end{pmatrix} \begin{pmatrix} \tilde{x} \\ \tilde{y} \\ \tilde{z} \end{pmatrix} = 9\tilde{x}^2 - 4\tilde{y}^2 - 36\tilde{z}^2
\end{aligned}$$

Verwenden wir noch $\begin{pmatrix} x \\ y \\ z \end{pmatrix} = U \begin{pmatrix} \tilde{x} \\ \tilde{y} \\ \tilde{z} \end{pmatrix}$, dann $8\sqrt{2}(y - z) = -16\sqrt{2}\tilde{y}$ und daher:

$$\begin{aligned}
-21x^2 - 5y^2 - 5z^2 + 30xy + 30xz - 2yz + 8\sqrt{2}y - 8\sqrt{2}z - 52 \\
&= 9\tilde{x}^2 - 4\tilde{y}^2 - 36\tilde{z}^2 - 16\tilde{y} - 52 \\
&= 9\tilde{x}^2 - 4(\tilde{y} + 2)^2 - 36\tilde{z}^2 - 36 \\
&= 36 \left\{ \left(\frac{\tilde{x}}{2}\right)^2 - \left(\frac{\tilde{y} + 2}{3}\right)^2 - \tilde{z}^2 - 1 \right\}
\end{aligned}$$

Somit ist $\alpha: \mathbb{R}^3 \rightarrow \mathbb{R}^3$,

$$\alpha \begin{pmatrix} x \\ y \\ z \end{pmatrix} = U^t \begin{pmatrix} x \\ y \\ z \end{pmatrix} + \begin{pmatrix} 0 \\ 2 \\ 0 \end{pmatrix} = \begin{pmatrix} \tilde{x} \\ \tilde{y} + 2 \\ \tilde{z} \end{pmatrix} = \begin{pmatrix} (x + y + z)/\sqrt{3} \\ (-y + z)/\sqrt{2} + 2 \\ (-2x + y + z)/\sqrt{6} \end{pmatrix}$$

eine Bewegung, sodass

$$\alpha(E) = \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3 : \left(\frac{x}{2}\right)^2 - \left(\frac{y}{3}\right)^2 - \left(\frac{z}{1}\right)^2 - 1 = 0 \right\}.$$

Insbesondere ist E ein zweischaliges Hyperboloid.

VIII.3. Projektive Räume und Abbildungen.

VIII.3.1. DEFINITION (Projektiver Raum). Sei V ein endlich dimensionaler Vektorraum über einem Körper \mathbb{K} . Unter dem *projektiven Raum* $P(V)$ verstehen wir die Menge aller Geraden durch den Ursprung in V , d.h.

$$P(V) := \{L \mid L \text{ is 1-dimensionaler linearer Teilraum von } V\}.$$

Jeder Vektor $0 \neq v \in V$ spannt einen 1-dimensionalen Teilraum $\langle v \rangle$ in V auf und repräsentiert daher einen Punkt in $\langle v \rangle \in P(V)$. Zwei Vektoren $v, v' \in V \setminus 0$ spannen genau dann den selben Teilraum auf, wenn ein Skalar $0 \neq \lambda \in \mathbb{K}$ existiert, sodass $v' = \lambda v$. Die Relation

$$v \sim v' \quad := \quad \exists \lambda \in \mathbb{K} \setminus 0 : v' = \lambda v$$

ist offensichtlich eine Äquivalenzrelation auf $V \setminus 0$. Die Menge der Äquivalenzklassen kann in kanonischer Weise mit $P(V)$ identifiziert werden,

$$P(V) = (V \setminus 0) / \sim, \quad \langle v \rangle \leftrightarrow [v].$$

Wir werden den von einem Vektor $0 \neq v \in V$ repräsentierten Punkt in $P(V)$ im folgenden mit $[v]$ bezeichnen.

Ist W ein linearer Teilraum von V , dann können wir jede Gerade in W auch als Gerade in V auffassen, und so $P(W)$ als Teilmenge von $P(V)$ betrachten. Für zwei Teilräume W und W' von V gilt offensichtlich:

$$P(W) = P(W') \quad \Leftrightarrow \quad W = W'. \quad (\text{VIII.14})$$

VIII.3.2. DEFINITION (Projektive Teilräume). Eine Teilmenge P von $P(V)$ wird als *projektiver Teilraum von* $P(V)$ bezeichnet, wenn sie von der Form $P = P(W)$ ist, wobei W ein linearer Teilraum von V ist. Die *Dimension eines projektiven Teilraums* $P = P(W)$ wird durch

$$\dim(P) := \dim(W) - 1$$

definiert. Da die Bedingung $P = P(W)$ den Teilraum W eindeutig festlegt, vgl. (VIII.14), ist dies wohldefiniert. 1-dimensionale projektive Teilräume heißen *projektive Geraden*, 2-dimensionale projektive Teilräume werden als *projektive Ebenen* bezeichnet. Gilt $\dim(P) = \dim(P(V)) - 1$, dann wird P *projektive Hyperbene* genannt.

VIII.3.3. BEISPIEL. Die leere Teilmenge bildet einen projektiven Teilraum von $P(V)$ und es gilt $\dim(\emptyset) = -1$. Die 0-dimensionalen projektiven Teilräume von $P(V)$ sind gerade die einpunktigen Teilmengen, d.h. jeder Punkt in $P(V)$, kann als 0-dimensionaler projektiver Teilraum aufgefasst werden. Auch $P = P(V)$ ist ein projektiver Teilraum von $P(V)$, es gilt $\dim(P(V)) = \dim(V) - 1$.

Sind $P \subseteq P' \subseteq P(V)$ zwei projektive Teilräume, so folgt $\dim(P) \leq \dim(P')$. Gilt Gleichheit, $\dim(P) = \dim(P')$, dann muss schon $P = P'$ sein. Ist nämlich $P = P(W)$ und $P' = P(W')$, so folgt $W \subseteq W'$ und $\dim(W) = \dim(W')$, also $W = W'$ und daher $P = P'$.

Der Durchschnitt zweier projektiver Teilräume W und W' von $P(V)$ ist ein projektiver Teilraum von $P(V)$, denn

$$P(W) \cap P(W') = P(W \cap W').$$

Allgemeiner bildet der Durchschnitt $\bigcap_i P_i$ beliebig vieler projektiver Teilräume, $P_i \subseteq P(V)$, einen projektiven Teilraum von $P(V)$. Sind nämlich W_i lineare Teilräume von V , sodass $P_i = P(W_i)$, dann gilt $\bigcap_i P_i = \bigcap_i P(W_i) = P(\bigcap_i W_i)$.

VIII.3.4. DEFINITION (Projektive Hülle). Die *projektive Hülle* einer Teilmenge $S \subseteq P(V)$ ist der (eindeutig bestimmte) kleinste projektive Teilraum, der S enthält. In anderen Worten,

$$\langle S \rangle_{\text{proj}} := \bigcap_{S \subseteq P} P,$$

wobei der Durchschnitt über alle projektiven Teilräume P von $P(V)$ genommen wird, die S enthalten.

Sind W und W' zwei lineare Teilräume von V , dann gilt

$$\langle P(W) \cup P(W') \rangle_{\text{proj}} = P(W + W'),$$

denn $W + W'$ ist der kleinste lineare Teilraum, der W und W' enthält.

VIII.3.5. PROPOSITION (Dimensionsformel). Sind P und P' zwei projektive Teilräume von $P(V)$, dann gilt

$$\dim(P) + \dim(P') = \dim(P \cap P') + \dim(\langle P \cup P' \rangle_{\text{proj}}).$$

BEWEIS. Nach Voraussetzung existieren lineare Teilräume W und W' von V , sodass $P = P(W)$ und $P' = P(W')$. Weiters ist $P \cap P' = P(W \cap W')$ und $\langle P \cup P' \rangle_{\text{proj}} = P(W + W')$. Es gilt daher:

$$\dim(P) = \dim(W) - 1$$

$$\dim(P') = \dim(W') - 1$$

$$\dim(P \cap P') = \dim(W \cap W') - 1$$

$$\dim(\langle P \cup P' \rangle_{\text{proj}}) = \dim(W + W') - 1$$

Die Dimensionsformel folgt somit aus der Dimensionsformel, siehe Satz IV.2.1,

$$\dim(W) + \dim(W') = \dim(W \cap W') + \dim(W + W')$$

für lineare Teilräume. □

Beachte, dass die analoge Dimensionsformel für affine Teilräume nur Fall nicht disjunkter Teilräume gültig ist, vgl. Beispiel VIII.1.24.

VIII.3.6. BEISPIEL. Zwei verschiedene Punkte (d.h. 0-dimensionale projektive Teilräume) $P \neq P' \in P(V)$ spannen eine projektive Gerade auf, denn aus der Dimensionsformel folgt

$$\dim(\langle P \cup P' \rangle_{\text{proj}}) = 1,$$

da ja $P \cap P' = \emptyset$, also $\dim(P \cap P') = -1$. Je zwei verschiedene Punkte in $P(V)$ liegen daher auf genau einer projektiven Geraden.

VIII.3.7. BEISPIEL. Sind P und P' zwei projektive Teilräume von $P(V)$ mit

$$\dim(P) + \dim(P') \geq \dim(P(V)),$$

so folgt aus der Dimensionsformel $\dim(P \cap P') \geq 0$ und daher $P \cap P' \neq \emptyset$. Insbesondere haben je zwei projektive Geraden in einer projektiven Ebene mindestens einen Schnittpunkt. Sind die beiden Geraden verschieden, so schneiden sie sich in genau einem Punkt, auch dies folgt aus der Dimensionsformel, denn die projektive Hülle zweier verschiedener Geraden in einer projektiven Ebene muss 2-dimensional sein. In der affinen Geometrie ist die Situation komplizierter, da hier Geraden auch parallel sein können.

VIII.3.8. BEISPIEL. Betrachte eine projektive Gerade G und eine projektive Ebene E in einem 3-dimensionalen projektiven Raum, $\dim(P(V)) = 3$. Aus der Dimensionsformel erhalten wir $\dim(E \cap G) \geq 0$, also $E \cap G \neq \emptyset$. Eine projektive Gerade und eine projektive Ebene im 3-dimensionalen projektiven Raum haben daher stets mindestens einen Schnittpunkt. Liegt die Gerade G nicht zur Gänze in E , dann schneiden sie sich in genau einem Punkt. In diesem Fall gilt nämlich $\dim(\langle G \cup E \rangle_{\text{proj}}) = 3$ und daher $\dim(E \cap G) = 0$. Auch hier ist die analoge affine Situation komplizierter, da die Gerade parallel zur Ebene sein kann.

VIII.3.9. BEISPIEL. Wir betrachten zwei projektive Ebenen E und E' in einem 3-dimensionalen projektiven Raum, $\dim(P(V)) = 3$. Aus der Dimensionsformel folgt $\dim(E \cap E') \geq 1$, d.h. der Durchschnitt $E \cap E'$ enthält eine projektive Gerade. Sind die beiden Ebenen verschieden, $E \neq E'$, dann ist ihr Durchschnitt $E \cap E'$ eine projektive Gerade, denn in diesem Fall gilt $\dim(\langle E \cup E' \rangle_{\text{proj}}) = 3$ und daher $\dim(E \cap E') = 1$. Wieder ist die affine Situation komplizierter, da Ebenen parallel sein können, dieser Fall tritt in der projektiven Geometrie nicht auf.

Jedes lineare Funktional, $\alpha: V \rightarrow \mathbb{K}$, definiert eine Teilmenge

$$\{[v] \in P(V) \mid \alpha(v) = 0\} = P(\ker(\alpha)).$$

Beachte, dass α keine Funktion auf $P(V)$ definiert, trotzdem ist die Nullstellenmenge wohldefiniert, denn aus $\alpha(v) = 0$ folgt $\alpha(\lambda v) = 0$, für jedes $\lambda \in \mathbb{K}$. Beachte, dass dies eine projektive Hyperebene bildet, falls $\alpha \neq 0$. Jeder projektive Teilraum P von $P(V)$ lässt sich in der Form

$$P = \{[v] \in P(V) \mid \alpha_1(v) = 0, \dots, \alpha_k(v) = 0\} = P\left(\bigcap_{i=1}^k \ker(\alpha_i)\right)$$

schreiben, für gewisse lineare Funktionale $\alpha_1, \dots, \alpha_k \in V^*$. Darüber hinaus kann $k = \dim(P(V)) - \dim(P)$ gewählt werden, und eine Darstellung mit weniger Funktionalen ist nicht möglich. Dies folgt daraus, dass jeder lineare Teilraum W von V in der Form $W = \bigcap_{i=1}^k \ker(\alpha_i)$ geschrieben werden kann, für geeignete Funktionale α_i , wobei $k = \dim(V) - \dim(W)$.

Wir wollen nun projektive und affine Geometrie in Zusammenhang bringen. Dazu wählen wir eine affine Hyperebene A in V , sodass $0 \notin A$. Jeder Punkt in A spannt daher einen 1-dimensionalen linearen Teilraum auf und definiert somit ein Element von $P(V)$. Dies liefert eine injektive Abbildung

$$\iota: A \rightarrow P(V), \quad \iota(v) := [v],$$

denn verschiedene Punkte in A spannen unterschiedliche 1-dimensionale Teilräume auf. Wir erhalten so aber nicht alle 1-dimensionalen Teilräume. Es gilt jedoch

$$P(V) = \iota(A) \sqcup P(V_A),$$

wobei V_A die zu A parallele lineare Hyperebene bezeichnet.¹⁴ Da $\dim(P(V_A)) = \dim(P(V)) - 1$, ist der affine Teil, $\iota(A)$, das Komplement einer projektiven Hyperebene. Beachte auch $\dim(A) = \dim(P(V))$. Der n -dimensionalen projektive Raum $P(V)$ zerfällt daher in einen n -dimensionalen affinen Raum A und eine $(n - 1)$ -dimensionale projektive Hyperebene, $P(V_A)$, deren Elemente als *Fernpunkte* bezeichnet werden. Ob ein Punkt in $P(V)$ Fernpunkt ist oder nicht, hängt von der Einbettung ι , d.h. von der Wahl des Teilraums A ab. Die Punkte in $P(V)$ sind alle gleichberechtigt, erst die Einbettung ι macht manche zu Fernpunkten. Beachte, dass jede projektive Hyperebene in $P(V)$ als Fernhyperebene einer geeigneten Einbettung ι auftreten kann.

VIII.3.10. BEISPIEL. Die Projektive Gerade zerfällt in eine affine Gerade und einen Fernpunkt. Die projektive Ebene zerfällt in eine affine Ebene und eine Ferngerade. Ein 3-dimensionaler projektiver Raum zerfällt in einen 3-dimensionalen affinen Raum und eine projektive Fernebene.

VIII.3.11. PROPOSITION. *Bezeichne $\iota: A \rightarrow P(V)$ die mit einer affinen Hyperebene A in V assoziierte Einbettung, $0 \notin A$. Dann gilt:*

(a) *Ist P' ein projektiver Teilraum von $P(V)$, so bildet $A' = \iota^{-1}(P')$ einen affinen Teilraum von A . Gilt $A' \neq \emptyset$, dann ist $\dim(A') = \dim(P')$ und*

$$P' = \iota(A') \sqcup P(V_{A'}),$$

wobei $V_{A'}$ den zu A' parallelen linearen Teilraum in V bezeichnet.

(b) *Ist A' ein affiner Teilraum von A , dann existiert ein projektiver Teilraum P' in $P(V)$, sodass $\iota^{-1}(P') = A'$. Gilt $A' \neq \emptyset$, dann ist P' eindeutig bestimmt, $P' = \langle \iota(A') \rangle_{proj} = P(\langle A' \rangle_{lin})$.*

¹⁴Das Zeichen \sqcup deutet eine *disjunkte Vereinigung* an, d.h. es ist als $P(V) = \iota(A) \cup P(V_A)$ und $\iota(A) \cap P(V_A) = \emptyset$ zu lesen.

BEWEIS. Ad (a): Nach Voraussetzung existiert ein linearer Teilraum W' von V , sodass $P' = P(W')$. Offensichtlich gilt $\iota^{-1}(P') = A \cap W'$, also bildet A' einen affinen Teilraum von A . Beachte, dass die affine Hülle von $A \cup W'$ echt größer als A ist, denn $0 \notin A$ aber $0 \in W'$. Da A eine affine Hyperebene bildet, folgt $\langle A \cup W' \rangle_{\text{aff}} = V$. Ist nun $A' \neq \emptyset$, dann erhalten wir aus der Dimensionsformel für affine Teilräume,

$$\underbrace{\dim(A)}_{=\dim(V)-1} + \dim(W') = \underbrace{\dim(A \cap W')}_{=\dim(A')} + \underbrace{\dim(\langle A \cup W' \rangle_{\text{aff}})}_{=\dim(V)},$$

also $\dim(A') = \dim(W') - 1 = \dim(P(W')) = \dim(P')$, siehe Beispiel VIII.1.24. In diesem Fall gilt weiters $P' \cap P(V_A) = P(W') \cap P(V_A) = P(W' \cap V_A) = P(V_{W' \cap A}) = P(V_{A'})$, also $P' = \iota(A') \sqcup P(V_{A'})$.

Ad (b): Sei A' ein affiner Teilraum von A und bezeichne $P' := \langle \iota(A') \rangle_{\text{proj}}$. Beachte $P' = P(W')$, wobei $W' = \langle A' \rangle_{\text{lin}}$. Somit $\iota^{-1}(P') = A \cap W' = A'$. Ist P'' ein weiterer projektiver Teilraum mit $\iota^{-1}(P'') = A'$, so muss $P' \subseteq P''$ gelten, denn P' ist der kleinste projektive Teilraum, der $\iota(A')$ enthält. Ist nun $A' \neq \emptyset$, dann erhalten wir aus der Dimensionsformel im ersten Teil $\dim(P') = \dim(P'')$ und daher $P' = P''$. Somit ist auch die Eindeutigkeit des Teilraums P' gezeigt. \square

VIII.3.12. BEISPIEL. Sei $\iota: A \rightarrow P(V)$ und $P(V) = \iota(A) \sqcup P(V_A)$ wie oben. Ist ein projektiver Teilraum P von $P(V)$ durch lineare Gleichungen gegeben,

$$P = \{[v] \in P(V) \mid \alpha_1(v) = 0, \dots, \alpha_k(v) = 0\}$$

wobei $\alpha_i \in V^*$, dann folgt

$$\iota^{-1}(P) = \{v \in A \mid \alpha_1(v) = 0, \dots, \alpha_k(v) = 0\}.$$

Die Einschränkungen $\alpha_i|_A: A \rightarrow \mathbb{K}$ liefern daher ein affines Gleichungssystem für den affinen Teilraum $\iota^{-1}(P)$.

Wir werden die Notation $\mathbb{K}P^n := P(\mathbb{K}^{n+1})$ verwenden, $\dim(\mathbb{K}P^n) = n$. Ist $0 \neq (x_0, \dots, x_n) \in \mathbb{K}^{n+1}$, dann bezeichnen wir den davon aufgespannten 1-dimensionalen linearen Teilraum mit $[x_0 : \dots : x_n] \in \mathbb{K}P^n$. Beachte, dass

$$[x_0 : \dots : x_n] = [x'_0 : \dots : x'_n] \iff \exists \lambda \in \mathbb{K} \setminus 0 : (x_0, \dots, x_n) = \lambda(x'_0, \dots, x'_n).$$

Die Skalare x_0, \dots, x_n werden als *homogene Koordinaten* bezeichnet, und sind nur bis auf ein gemeinsames Vielfaches bestimmt. Die mit der affinen Hyperebene $A = e_0 + \langle e_1, \dots, e_n \rangle_{\text{lin}}$ assoziierte Einbettung,

$$\iota: \mathbb{K}^n \rightarrow \mathbb{K}P^n, \quad \iota(x_1, \dots, x_n) = [1 : x_1 : \dots : x_n]$$

wird als *Standardeinbettung* des affinen Raums bezeichnet. Wird erhalten die Zerlegung

$$\mathbb{K}P^n = \iota(\mathbb{K}^n) \sqcup \mathbb{K}P^{n-1},$$

wobei die Einbettung $\mathbb{K}P^{n-1} \subseteq \mathbb{K}P^n$ durch $[x_1 : \cdots : x_n] \mapsto [0 : x_1 : \cdots : x_n]$ gegeben ist. Ist eine projektive Hyperebene durch eine Gleichung gegeben,

$$P = \{[x_0 : \cdots : x_n] \in \mathbb{K}P^n \mid a_0x_0 + a_1x_1 + \cdots + a_nx_n = 0\}$$

dann erhalten wir sofort eine affine Gleichung für den affinen Teil,

$$\iota^{-1}(P) = \{(x_1, \dots, x_n) \in \mathbb{K}^n \mid a_0 + a_1x_1 + \cdots + a_nx_n = 0\}.$$

Analoges gilt für projektive Teilräume beliebiger Dimension.

VIII.3.13. BEISPIEL. Wir betrachten zwei Punkte $p_1 = (\frac{1}{2})$ und $p_2 = (\frac{3}{4})$ in \mathbb{R}^2 . Die Gerade durch p_1 und p_2 ist daher

$$g = \langle p_1, p_2 \rangle_{\text{aff}} = \left\{ \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \in \mathbb{R}^2 : 1 + x_1 - x_2 = 0 \right\}.$$

Mit Hilfe der Standardeinbettung erhalten wir zwei Punkte $P_1 = \iota(p_1) = [1 : 1 : 2]$ und $P_2 = \iota(p_2) = [1 : 3 : 4]$ in der projektiven Ebene $\mathbb{R}P^2$. Die projektive Gerade durch P_1 und P_2 ist

$$G = \langle P_1, P_2 \rangle_{\text{proj}} = \{[x_0 : x_1 : x_2] \in \mathbb{R}P^2 : x_0 + x_1 - x_2 = 0\}.$$

Der affine Teil von G stimmt mit g überein, $\iota^{-1}(G) = g$, die projektive Gerade G besitzt aber auch einen Fernpunkt, $G \cap \mathbb{R}P^1 = \{[0 : 1 : 1]\}$, die Richtung der Geraden g . Sei nun

$$g' = \left\{ \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \in \mathbb{R}^2 : 2 + x_1 - x_2 = 0 \right\}$$

eine weitere affine Gerade. Beachte, $g \cap g' = \emptyset$, die beiden Geraden sind parallel. Bezeichnet

$$G' = \{[x_0 : x_1 : x_2] : 2x_0 + x_1 - x_2 = 0\}$$

die entsprechende projektive Gerade, $g' = \iota^{-1}(G')$, dann gilt jedoch

$$G \cap G' = \{[0 : 1 : 1]\}.$$

In diesem Sinn haben g und g' einen Schnittpunkt im Fernen, d.h. auf der Ferngeraden. Betrachten wir eine andere Einbettung, etwa

$$\tilde{\iota}: \mathbb{R}^2 \rightarrow \mathbb{R}P^2, \quad \tilde{\iota} \begin{pmatrix} x_0 \\ x_2 \end{pmatrix} = [x_0 : 1 : x_2],$$

dann sind die entsprechenden affinen Geraden durch andere Gleichungen gegeben:

$$\begin{aligned} \tilde{g} &:= \tilde{\iota}^{-1}(G) = \left\{ \begin{pmatrix} x_0 \\ x_2 \end{pmatrix} : x_0 + 1 - x_2 = 0 \right\}, \\ \tilde{g}' &:= \tilde{\iota}^{-1}(G') = \left\{ \begin{pmatrix} x_0 \\ x_2 \end{pmatrix} : 2x_0 + 1 - x_2 = 0 \right\}, \end{aligned}$$

ihr Schnittpunkt $\tilde{g} \cap \tilde{g}' = \left\{ \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\}$ liegt nun im affinen Teil.

VIII.3.14. BEISPIEL. Wir betrachten die affine Ebene

$$\varepsilon = \left\{ \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \in \mathbb{R}^3 : 2 + 3x_1 + 4x_2 + 5x_3 = 0 \right\}.$$

Die entsprechende projektive Ebene ist daher durch

$$E = \{[x_0 : x_1 : x_2 : x_3] \in \mathbb{R}P^3 \mid 2x_0 + 3x_1 + 4x_2 + 5x_3 = 0\}$$

gegeben, $\iota^{-1}(E) = \varepsilon$. Für die Fernpunkte von E erhalten wir

$$E \cap \mathbb{RP}^2 = \{[0 : x_1 : x_2 : x_3] \in \mathbb{RP}^3 : 3x_1 + 4x_2 + 5x_3 = 0\}.$$

Die projektive Ebene zerfällt daher in die affine Ebene ε und eine projektive Gerade im Fernen. Betrachte nun die Gerade $g = \langle \left(\frac{6}{7}\right), \left(\frac{7}{9}\right) \rangle_{\text{aff}}$ in \mathbb{R}^3 und bezeichne $G \subseteq \mathbb{RP}^3$ die entsprechende projektive Gerade, $\iota^{-1}(G) = g$. Die projektive Gerade G besteht aus der affinen Geraden g und einem Fernpunkt,

$$G \cap \mathbb{RP}^2 = \{[0 : 1 : -2 : 1]\}.$$

Beachte, dass ε und g parallel sind, $\varepsilon \cap g = \emptyset$. Es gilt jedoch

$$E \cap G = \{[0 : 1 : -2 : 1]\},$$

es gibt daher einen Schnittpunkt auf der Fernebene.

Jede injektive lineare Abbildung, $\phi: V \rightarrow V'$, induziert eine Abbildung

$$P_\phi: P(V) \rightarrow P(V'), \quad P_\phi([v]) := [\phi(v)].$$

Beachte, dass dies tatsächlich wohldefiniert ist. Offensichtlich gilt

$$P_{\text{id}_V} = \text{id}_{P(V)} \quad \text{und} \quad P_{\phi' \circ \phi} = P_{\phi'} \circ P_\phi,$$

für jede weitere injektive lineare Abbildung $\phi': V' \rightarrow V''$.

VIII.3.15. DEFINITION (Projektive Abbildungen). Seien V und V' zwei Vektorräume gleicher Dimension. Eine Abbildung $\pi: P(V) \rightarrow P(V')$ wird *projektiv* genannt, wenn sie von der Form $\pi = P_\phi$ ist, wobei $\phi: V \rightarrow V'$ ein linearer Isomorphismus ist. Unter einer *Projektivität* verstehen wir eine projektive Abbildung $\pi: P(V) \rightarrow P(V)$. Die Menge aller Projektivitäten bezeichnen wir mit $\text{Pro}(P(V))$ oder $\text{PGL}(V)$.

Beachte, dass projektive Abbildungen invertierbar sind, $(P_\phi)^{-1} = P_{\phi^{-1}}$. Insbesondere bildet die Menge der Projektivitäten bezüglich der Komposition von Abbildungen eine Gruppe.

VIII.3.16. LEMMA. Die Abbildung $\text{GL}(V) \rightarrow \text{Pro}(P(V))$, $\phi \mapsto P_\phi$, ist ein surjektiver Gruppenhomomorphismus mit Kern

$$\{\phi \in \text{GL}(V) : P_\phi = \text{id}_{P(V)}\} = \{\lambda \text{id}_V : \lambda \in \mathbb{K} \setminus \{0\}\}.$$

BEWEIS. Die Inklusion \supseteq offensichtlich. Für die umgekehrte Inklusion sei nun $\phi \in \text{GL}(V)$ so, dass $P_\phi = \text{id}_{P(V)}$. Es existieren daher Skalare $\lambda_v \in \mathbb{K}$, sodass $\phi(v) = \lambda_v v$. Aus der Linearität erhalten wir $\lambda_{\mu v} \mu v = \phi(\mu v) = \mu \phi(v) = \mu \lambda_v v$ und daher $\lambda_{\mu v} = \lambda_v$, für alle $0 \neq v \in V$ und $0 \neq \mu \in \mathbb{K}$. Weiters:

$$\lambda_{v_1+v_2}(v_1+v_2) = \phi(v_1+v_2) = \phi(v_1) + \phi(v_2) = \lambda_{v_1}v_1 + \lambda_{v_2}v_2. \quad (\text{VIII.15})$$

Sind v_1 und v_2 linear unabhängig in V , dann folgt $\lambda_{v_1} = \lambda_{v_1+v_2} = \lambda_{v_2}$. Daraus lässt sich nun leicht schließen, dass die Skalare λ_v alle übereinstimmen. Es existiert daher $\lambda \in \mathbb{K}$, sodass $\phi(v) = \lambda v$, für alle $v \in V$. Somit gilt $\phi = \lambda \text{id}_V$, womit auch die Inklusion \subseteq gezeigt wäre. \square

VIII.3.17. PROPOSITION. Sind V und V' zwei Vektorräume gleicher Dimension, dann gilt:

- (a) Ist P ein projektive Teilraum von $P(V)$ und $\pi: P(V) \rightarrow P(V')$ eine projektive Abbildung, dann bildet $\pi(P)$ einen projektiven Teilraum von $P(V')$ und es gilt $\dim(\pi(P)) = \dim(P)$.
- (b) Sind P und P' projektive Teilräume gleicher Dimension in $P(V)$ bzw. P' dann existiert eine projektive Abbildung $\pi: P(V) \rightarrow P(V')$, sodass $\pi(P) = P'$.

BEWEIS. Sei $P = P(W)$ ein projektiver Teilraum von $P(V)$, wobei W einen linearen Teilraum von V bezeichnet. Weiters sei $\pi = P_\phi: P(V) \rightarrow P(V')$ eine projektive Abbildung., wobei $\phi: V \rightarrow V'$ eine invertierbare lineare Abbildung bezeichnet. Aufgrund der Linearität von ϕ gilt $\pi(P) = P(\phi(W))$. Also ist $\pi(P)$ ein projektiver Teilraum von P' . Da ϕ injektiv ist haben wir auch $\dim(\phi(W)) = \dim(W)$, also $\dim(\pi(P)) = \dim(P)$. Damit ist (a) gezeigt.

Um (b) einzusehen, seien nun W und W' lineare Teilräume in V und V' , sodass $P = P(W)$ und $P' = P(W')$. Nach Voraussetzung gilt $\dim(W) = \dim(W')$. Es existiert daher eine invertierbare lineare Abbildung $\phi: V \rightarrow V'$, sodass $\phi(W) = W'$. Für die projektive Abbildung $\pi = P_\phi$ gilt dann $\pi(P) = P'$. \square

VIII.3.18. BEMERKUNG. Nach Proposition VIII.3.17(b) existiert eine projektive Abbildung $P(V) \rightarrow \mathbb{K}P^n$, wobei $n = \dim(P(V))$. Jeder projektive Raum ist daher zu $\mathbb{K}P^n$ projektiv äquivalent.

VIII.3.19. PROPOSITION. Sei $A \subseteq V$ eine affine Hyperebene, $0 \notin A$ und $\iota: A \rightarrow P(V)$ die damit assoziierte Einbettung, $P(V) = \iota(A) \sqcup P(V_A)$. Zu jedem affinen Isomorphismus $\alpha: A \rightarrow A$ existiert eine eindeutige Projektivität $\pi_\alpha: P(V) \rightarrow P(V)$, sodass $\pi_\alpha \circ \iota = \iota \circ \alpha$.¹⁵ Darüber hinaus stimmt die Einschränkung $\pi_\alpha|_{P(V_A)}: P(V_A) \rightarrow P(V_A)$ mit P_{ϕ_α} überein, wobei $\phi_\alpha: V_A \rightarrow V_A$ den linearen Teil von α bezeichnet. Wir erhalten so einen injektiven Gruppenhomomorphismus

$$\text{Aff}(A) \rightarrow \text{Pro}(P(V)), \quad \alpha \mapsto \pi_\alpha,$$

dessen Bild aus genau jenen Projektivitäten besteht, die $\iota(A)$ auf $\iota(A)$, oder äquivalent $P(V_A)$ auf $P(V_A)$, abbilden.

BEWEIS. Sei $\alpha: A \rightarrow A$ ein affiner Isomorphismus. Dann existiert ein eindeutiger affiner Isomorphismus $\phi: V \rightarrow V$, sodass $\phi|_A = \alpha$ und $\phi(0) = 0$, siehe Beispiel VIII.1.30. Aufgrund der zweiten Bedingung ist ϕ linear. Somit ist $\pi_\alpha := P_\phi$ eine Projektivität für die $\pi_\alpha \circ \iota = \iota \circ \alpha$ gilt. Offensichtlich gilt für den linearen Teil, $\phi_\alpha: V_A \rightarrow V_A$, von α auch $\phi_\alpha = \phi|_{V_A}$ und daher $\pi_\alpha|_{P(V_A)} = P_\phi|_{P(V_A)} = P_{\phi|_{V_A}} = P_{\phi_\alpha}: P(V_A) \rightarrow P(V_A)$. Darüber hinaus gilt $\pi_{\alpha' \circ \alpha} = \pi_{\alpha'} \circ \pi_\alpha$, also ist die Zuordnung $\alpha \mapsto \pi_\alpha$ ein Gruppenhomomorphismus und offensichtlich auch injektiv.

¹⁵Unterdrücken wir die Einbettung $\iota: A \rightarrow P(V)$ und fassen A als Teilmenge von $P(V)$ auf, dann bedeutet diese Gleichung gerade $\pi_\alpha|_A = \alpha$, d.h. π_α ist eine (projektive) Fortsetzung der affinen Abbildung α .

Schließlich ist noch zu zeigen, dass jede Projektivität $\pi: P(V) \rightarrow P(V)$, die $P(V_A)$ auf $P(V_A)$ abbildet von der Form $\pi = \pi_\alpha$ ist, wobei $\alpha \in \text{Aff}(A)$. Nach Voraussetzung ist $\pi = P_\phi$, wobei $\phi \in \text{GL}(V)$ und $\phi(V_A) = V_A$. Wir fixieren einen Punkt $a \in A$. Da $\phi(a) \notin V_A$ existiert $\lambda \in \mathbb{K}$, sodass $\lambda\phi(a) \in A$. O.B.d.A. dürfen wir daher $\phi(a) \in A$ annehmen. Daraus folgt $\phi(A) = \phi(a + V_A) = \phi(a) + \phi(V_A) = \phi(a) + V_A = A$. Wir schließen, dass die Einschränkung von ϕ einen affinen Isomorphismus $\alpha := \phi|_A: A \rightarrow A$ definiert. Nach Konstruktion gilt $\pi_\alpha = P_\phi = \pi$.

Lässt die Projektivität π alle Punkte in $\iota(A)$ fest, d.h. $\pi \circ \iota = \iota$, dann folgt aus obigen Betrachtungen sofort $\pi = P_{\text{id}_V} = \text{id}_{P(V)}$. Dies zeigt, dass die Projektivität π durch die Gleichung $\pi \circ \iota = \iota \circ \alpha$ eindeutig bestimmt ist. \square

VIII.3.20. BEISPIEL. Seien $\iota: A \rightarrow P(V)$ und $P(V) = \iota(A) \sqcup P(V_A)$ wie oben. Betrachte eine Translation $\tau \in \text{Aff}(A)$. Die Projektivität $\pi_\tau: P(V) \rightarrow P(V)$ lässt dann alle Punkte der Fernhyperebene $P(V_A)$ fest. Dies folgt aus Proposition VIII.3.19, denn für den linearen Teil einer Translation gilt $\phi_\tau = \text{id}_{V_A}$. Umgekehrt muss eine Projektivität, die alle Punkte der Fernhyperebene festlässt, eine affine Abbildung mit linearem Teil λid_{V_A} sein, siehe Lemma VIII.3.16. Die Projektivität, die die Fernhyperebene punktweise festlassen sind daher genau die Kompositionen von Translationen und Streckungen.

Ist $\alpha: \mathbb{K}^n \rightarrow \mathbb{K}^n$ ein affiner Isomorphismus, $\alpha(x) = Ax + b$, wobei $A \in \text{GL}_n(\mathbb{K})$ und $b \in \mathbb{K}^n$, dann wird die entsprechende Projektivität $\pi_\alpha: \mathbb{K}P^n \rightarrow \mathbb{K}P^n$ durch die Matrix $\begin{pmatrix} 1 & 0 \\ b & A \end{pmatrix} \in \text{GL}_{n+1}(\mathbb{K})$ repräsentiert. Die auf der Fernhyperebene induzierte Projektivität, $\pi_\alpha|_{\mathbb{K}P^{n-1}}: \mathbb{K}P^{n-1} \rightarrow \mathbb{K}P^{n-1}$, ist durch die Matrix A gegeben.

Sei nun $\pi: \mathbb{K}P^n \rightarrow \mathbb{K}P^n$ eine beliebige Projektivität und

$$\begin{pmatrix} d & c^t \\ b & A \end{pmatrix} \in \text{GL}_{n+1}(\mathbb{K})$$

eine Matrix die π darstellt, $d \in \mathbb{K}$, $b, c \in \mathbb{K}^n$, $A \in M_{n \times n}(\mathbb{K})$. Die nur teilweise definierte Einschränkung auf den affinen Teil, $\mathbb{K}^n \subseteq \mathbb{K}P^n$, ist durch eine rationale Funktion gegeben,

$$x \mapsto \frac{1}{c^t x + d}(Ax + b), \quad x \in \mathbb{K}^n.$$

Die affine Hyperebene auf der der Nenner verschwindet, $\{x \in \mathbb{K}^n : c^t x + d = 0\}$, besteht genau aus jenen Punkten die durch π auf die Fernhyperebene abgebildet werden. Insbesondere sind die Projektivitäten $\pi: \mathbb{K}P^1 \rightarrow \mathbb{K}P^1$ auf dem affinen Teil $\mathbb{K} \subseteq \mathbb{K}P^1$ genau die rationalen Abbildungen der Form

$$x \mapsto \frac{ax + b}{cx + d}, \quad x \in \mathbb{K}.$$

wobei $ad - bc \neq 0$.

VIII.3.21. BEISPIEL. Die Matrix $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ stellt eine Projektivität $\pi: \mathbb{K}P^1 \rightarrow \mathbb{K}P^1$ dar. Diese lässt den Punkt $\iota(1) = [1 : 1]$ fest und vertauscht $\iota(0) = [1 : 0]$ mit dem Fernpunkt $[0 : 1]$. Auf dem affinen Teil ist sie durch $x \mapsto x^{-1}$ gegeben.

Ist $\mathbb{K} = \mathbb{R}$ und $\dim(P(V)) \geq 2$, dann können die Projektivitäten auf $P(V)$ auch als jene Bijektionen charakterisiert werden, die projektive Geraden auf projektive Geraden abbilden. Genauer haben wir:

VIII.3.22. SATZ (Fundamentalsatz der projektiven Geometrie). *Sei V ein endlich dimensionaler reeller Vektorraum und $\dim(P(V)) \geq 2$. Dann ist jede projektive Kollineation, d.h. jede Bijektion $\pi: P(V) \rightarrow P(V)$, die projektive Geraden auf projektive Geraden abbildet, eine Projektivität.*

BEWEIS. Wir werden dies auf den Fundamentalsatz der affinen Geometrie zurückführen, siehe Satz VIII.1.32. Wir zeigen zunächst, dass π auch k -dimensionale projektive Teilräume auf k -dimensionale projektive Teilräume abbildet. Wir zeigen dies mittels Induktion nach k . Die Fälle $k = -1$ und $k = 0$ sind trivial.¹⁶ Für den Induktionsschritt sei nun P ein k -dimensionaler projektiver Teilraum von $P(V)$ und $k \geq 1$. Wähle einen $(k-1)$ -dimensionalen projektiven Teilraum $P' \subseteq P$ und einen Punkt $p \in P \setminus P'$. Jede projektive Gerade in P , die p enthält, muss P' schneiden, siehe Beispiel VIII.3.7. Es gilt daher

$$P = \bigcup_{p' \in P'} \langle p, p' \rangle_{\text{proj}}.$$

Da π projektive Geraden auf projektive Geraden abbildet, folgt

$$\pi(P) = \bigcup_{p' \in P'} \pi(\langle p, p' \rangle_{\text{proj}}) = \bigcup_{p' \in P'} \langle \pi(p), \pi(p') \rangle_{\text{proj}} = \bigcup_{\tilde{p}' \in \pi(P')} \langle \pi(p), \tilde{p}' \rangle_{\text{proj}}.$$

Nach Induktionsvoraussetzung bildet $\pi(P')$ einen $(k-1)$ -dimensionalen projektiven Teilraum von $P(V)$. Da $\pi(p) \notin \pi(P')$ folgt aus obiger Darstellung, dass $\pi(P)$ einen k -dimensionalen projektiven Teilraum bildet.

Sei nun $A \subseteq V$ eine affine Hyperebene, $0 \notin A$ und $P(V) = \iota(A) \sqcup P(V_A)$ die damit assoziierte Zerlegung. Nach dem vorangehenden Absatz ist $\pi(P(V_A))$ eine projektive Hyperebene. Nach Proposition VIII.3.17(b) existiert eine Projektivität $\pi': P(V) \rightarrow P(V)$, sodass $\pi'(\pi(P(V_A))) = P(V_A)$. O.B.d.A. dürfen wir daher annehmen, dass π die Fernhyperebene bewahrt, d.h. $\pi(P(V_A)) = P(V_A)$. Es existiert daher eine Bijektion $\alpha: A \rightarrow A$, sodass $\pi \circ \iota = \iota \circ \alpha$. Da π projektive Geraden auf projektive Geraden abbildet, muss α affine Geraden auf affine Geraden abbilden, siehe Proposition VIII.3.11. Nach dem Fundamentalsatz der affinen Geometrie ist α daher ein affiner Isomorphismus, siehe Satz VIII.1.32. Nach Proposition VIII.3.19 existiert eine Projektivität $\pi'': P(V) \rightarrow P(V)$, sodass $\pi'' \circ \iota = \iota \circ \alpha$. O.B.d.A. dürfen wir daher $\pi|_{\iota(A)} = \text{id}_{\iota(A)}$ annehmen. Jeder Punkt in $P(V_A)$ lässt sich als Schnittpunkt zweier verschiedener projektiver Geraden schreiben, die nicht zur Gänze in $P(V_A)$ liegen. Da π projektive Geraden auf projektive Geraden abbildet, und weil $\pi|_{\iota(A)} = \text{id}_{\iota(A)}$, muss π auch alle Punkte in $P(V_A)$ fest lassen, d.h. $\pi|_{P(V_A)} = \text{id}|_{P(V_A)}$. Somit $\pi = \text{id}_{P(V)}$. \square

¹⁶Der Fall $k = 1$ ist unsere Voraussetzung an π .

Wir betrachten wieder einen allgemeinen Körper \mathbb{K} .

Eine projektive Abbildung $\pi: P(V) \rightarrow P(V')$ ist durch die Bilder von $n + 1$ Punkten noch nicht eindeutig bestimmt, $n = \dim(P(V)) = \dim(P(V'))$. Etwa liefert die Matrix $\begin{pmatrix} 1 & 0 \\ 0 & \lambda \end{pmatrix}$, wobei $0 \neq \lambda \in \mathbb{K}$, eine Projektivität $\pi: \mathbb{K}P^1 \rightarrow \mathbb{K}P^1$, mit

$$\pi([1 : 0]) = [1 : 0], \quad \pi([0 : 1]) = [0 : 1], \quad \pi([1 : 1]) = [1 : \lambda].$$

Ist darüber hinaus $\lambda \neq 1$, dann gilt $\pi([1 : 1]) \neq [1 : 1]$, also $\pi \neq \text{id}_{\mathbb{K}P^1}$. Die Einschränkung auf den affinen Teil ist in diesem Fall eine Streckung, $x \mapsto \lambda x$.

VIII.3.23. DEFINITION (Projektive Bezugssysteme). Sei $P(V)$ ein projektiver Raum und $n = \dim(P(V))$. Punkte $P_0, \dots, P_{n+1} \in P(V)$ werden als *projektives Bezugssystem* bezeichnet, falls die projektive Hülle von je $n + 1$ dieser Punkte mit $P(V)$ übereinstimmt. In diesem Fall sagen wir auch die Punkte P_0, \dots, P_{n+1} sind in *allgemeiner Lage*.

VIII.3.24. LEMMA. Seien $v_0, \dots, v_{n+1} \in V \setminus 0$, wobei $n = \dim(P(V))$. Dann sind äquivalent:

- (a) Die Punkte $[v_0], \dots, [v_{n+1}]$ bilden ein projektives Bezugssystem von $P(V)$.
- (b) Je $n + 1$ der Vektoren v_0, \dots, v_{n+1} bilden eine Basis von V .
- (c) v_0, \dots, v_n ist eine Basis von V und in der Darstellung $v_{n+1} = \sum_{i=0}^n \mu_i v_i$ sind alle Koeffizienten verschieden von Null, d.h. $\mu_i \neq 0$ für alle, $i = 0, \dots, n$.

BEWEIS. Sind $w_0, \dots, w_n \in V \setminus 0$, dann gilt

$$\langle [w_0], \dots, [w_n] \rangle_{\text{proj}} = P(\langle w_0, \dots, w_n \rangle_{\text{lin}}).$$

Es gilt daher $\langle [w_0], \dots, [w_n] \rangle_{\text{proj}} = P(V)$ genau dann wenn w_0, \dots, w_n eine Basis von V bilden. Daraus folgt die Äquivalenz der ersten beiden Aussagen.

In der Darstellung $v_{n+1} = \sum_{i=0}^n \mu_i v_i$ müssen alle $\mu_i \neq 0$ sein, denn andernfalls wären die restlichen $n + 1$ Vektoren v_i linear abhängig. Es gilt daher (b) \Rightarrow (c). Die Implikation (c) \Rightarrow (b) folgt daraus, dass je $n + 1$ der Vektoren v_0, \dots, v_{n+1} ein Erzeugendensystem von V bilden. \square

VIII.3.25. BEISPIEL. Die Punkte

$$P_0 = [1 : 0 : \dots : 0], \dots, P_n = [0 : \dots : 0 : 1], P_{n+1} = [1 : 1 : \dots : 1]$$

bilden ein projektives Bezugssystem von $\mathbb{K}P^n$.

VIII.3.26. BEISPIEL. Drei Punkte einer projektiven Gerade bilden genau dann ein projektives Bezugssystem, wenn sie paarweise verschieden sind.

VIII.3.27. PROPOSITION. Ist $\dim(P(V)) = n = \dim(P(V'))$ dann gilt:

- (a) Bilden die Punkte P_0, \dots, P_{n+1} ein projektives Bezugssystem von $P(V)$ und ist $\pi: P(V) \rightarrow P(V')$ eine projektive Abbildung, dann bilden auch die Punkte $\pi(P_0), \dots, \pi(P_{n+1})$ ein projektives Bezugssystem von $P(V')$.
- (b) Sind P_0, \dots, P_{n+1} und P'_0, \dots, P'_{n+1} projektive Bezugssysteme von $P(V)$ bzw. $P(V')$, dann existiert eine eindeutige projektive Abbildung $\pi: P(V) \rightarrow P(V')$, sodass $\pi(P_i) = P'_i$, für alle $i = 0, \dots, n + 1$.

BEWEIS. Die erste Aussage ist offensichtlich.

Wir werden nun die Existenz einer projektiven Abbildung $\pi: P(V) \rightarrow P(V')$ mit $\pi(P_i) = P'_i$ zeigen. Wir wählen dazu Vektoren $0 \neq v_i \in V$ und $0 \neq v'_i \in V'$, sodass $P_i = [v_i]$ und $P'_i = [v'_i]$, $i = 0, \dots, n+1$. Nach Lemma VIII.3.24 bilden v_0, \dots, v_n und v'_0, \dots, v'_n Basen von V bzw. V' und existieren Skalare $0 \neq \mu_i \in \mathbb{K}$ und $0 \neq \mu'_i \in \mathbb{K}$, $i = 0, \dots, n$, sodass

$$v_{n+1} = \sum_{i=0}^n \mu_i v_i \quad \text{und} \quad v'_{n+1} = \sum_{i=0}^n \mu'_i v'_i.$$

Es bezeichne $\phi: V \rightarrow V'$ die eindeutige lineare Abbildung mit

$$\phi(v_i) = \frac{\mu'_i}{\mu_i} v'_i, \quad i = 0, \dots, n.$$

Da ϕ die Basis v_0, \dots, v_n auf eine Basis von V' abbildet, ist ϕ invertierbar. Nach Konstruktion gilt weiters

$$\phi(v_{n+1}) = \phi\left(\sum_{i=0}^n \mu_i v_i\right) = \sum_{i=0}^n \mu_i \phi(v_i) = \sum_{i=0}^n \mu_i \mu'_i v'_i = v'_{n+1}.$$

Die damit assoziierte projektive Abbildung, $\pi = P_\phi: P(V) \rightarrow P(V')$, erfüllt daher $\pi(P_i) = P'_i$, für alle $i = 0, \dots, n+1$.

Schließlich widmen wir uns der Eindeutigkeit der projektiven Abbildung. Es genügt zu zeigen, dass eine Projektivität $\pi: P(V) \rightarrow P(V)$, die die Punkte P_i alle festlässt, d.h. $\pi(P_i) = P_i$, $i = 0, \dots, n+1$, schon die identische Abbildung sein muss, $\pi = \text{id}_{P(V)}$. Wie zuvor seien $0 \neq v_i \in V$, sodass $P_i = [v_i]$, $i = 0, \dots, n+1$ und $0 \neq \mu_i \in \mathbb{K}$ mit $v_{n+1} = \sum_{i=0}^n \mu_i v_i$. Weiters sei $\phi \in \text{GL}(V)$, sodass $\pi = P_\phi$. Nach Voraussetzung existieren daher $0 \neq \lambda_i \in \mathbb{K}$, sodass $\phi(v_i) = \lambda_i v_i$, für $i = 0, \dots, n+1$. Aus der Linearität von ϕ erhalten wir

$$\sum_{i=0}^n \lambda_{n+1} \mu_i v_i = \lambda_{n+1} v_{n+1} = \phi(v_{n+1}) = \phi\left(\sum_{i=0}^n \mu_i v_i\right) = \sum_{i=0}^n \mu_i \lambda_i v_i.$$

Da v_0, \dots, v_n eine Basis von V bildet folgt $\mu_i \lambda_{n+1} = \mu_i \lambda_i$ und dann $\lambda_{n+1} = \lambda_i$, für alle $i = 0, \dots, n$, denn $\mu_i \neq 0$. Somit stimmen alle λ_i überein, also $\phi = \lambda \text{id}_V$ und daher $\pi = P_\phi = \text{id}_{P(V)}$. \square

VIII.3.28. BEISPIEL. Ist P_0, \dots, P_{n+1} ein projektives Bezugssystem von $P(V)$, dann existiert eine eindeutige projektive Abbildung $\pi: \mathbb{K}P^n \rightarrow P(V)$, sodass:

$$\begin{aligned} \pi([1 : 0 : \dots : 0]) &= P_0 \\ &\vdots \\ \pi([0 : \dots : 0 : 1]) &= P_n \\ \pi([1 : 1 : \dots : 1]) &= P_{n+1}. \end{aligned}$$

VIII.3.29. DEFINITION (Doppelverhältnis). Seien P_0, P_1, P_2, P_3 vier Punkte auf einer projektiven Gerade G , sodass P_0, P_1, P_2 paarweise verschieden sind. Weiters sei $\pi: \mathbb{K}P^1 \rightarrow G$ die eindeutige projektive Abbildung mit

$$\pi([1 : 0]) = P_0, \quad \pi([0 : 1]) = P_1 \quad \text{und} \quad \pi([1 : 1]) = P_2.$$

Bezeichnet $[x_0 : x_1] \in \mathbb{K}P^1$ jenen Punkt mit $\pi([x_0 : x_1]) = P_3$, dann wird

$$DV(P_0, P_1, P_2, P_3) := \frac{x_1}{x_0} \in \mathbb{K} \cup \{\infty\}$$

als Doppelverhältnis der Punkte P_0, P_1, P_2, P_3 bezeichnet. Beachte, dass dieser Quotient wohldefiniert ist, denn x_0 und x_1 sind bis auf ein gemeinsames Vielfaches bestimmt. Falls $x_0 = 0$, dann ist dies als $\frac{x_1}{x_0} = \infty$ zu verstehen.

VIII.3.30. BEISPIEL. Vermöge der Einbettung $\mathbb{K} \subseteq \mathbb{K}P^1$, $x \mapsto [1 : x]$, ist

$$\mathbb{K}P^1 = \mathbb{K} \cup \{\infty\},$$

wobei $\infty := [0 : 1] \in \mathbb{K}P^1$ den Fernpunkt bezeichnet. Für jedes $x \in \mathbb{K} \cup \{\infty\}$ gilt

$$DV(0, \infty, 1, x) = x.$$

VIII.3.31. BEISPIEL. Seien P_0, P_1, P_2 drei paarweise verschiedene Punkte einer projektiven Gerade. Dann gilt:

$$DV(P_0, P_1, P_2, P_0) = 0, \quad DV(P_0, P_1, P_2, P_1) = \infty, \quad DV(P_0, P_1, P_2, P_2) = 1.$$

VIII.3.32. BEISPIEL. Seien P_0, P_1, P_2, P_3 vier paarweise verschiedene Punkte einer projektiven Gerade. Dann gilt:

$$DV(P_1, P_0, P_2, P_3) = \frac{1}{DV(P_0, P_1, P_2, P_3)} = DV(P_0, P_1, P_3, P_2)$$

sowie

$$DV(P_2, P_1, P_0, P_3) = 1 - DV(P_0, P_1, P_2, P_3) = DV(P_0, P_3, P_2, P_1)$$

siehe Aufgabe 46. Da diese Transpositionen die Gruppe \mathfrak{S}_4 erzeugen, lässt sich damit $DV(P_{\sigma(0)}, P_{\sigma(1)}, P_{\sigma(2)}, P_{\sigma(3)})$ für jede Permutation σ aus $DV(P_0, P_1, P_2, P_3)$ berechnen.

Projektive Abbildungen bewahren das Doppelverhältnis:

VIII.3.33. PROPOSITION. *Sei $\pi : P(V) \rightarrow P(V')$ eine projektive Abbildung. Weiters seien P_0, P_1, P_2, P_3 vier Punkt auf einer projektiven Gerade in $P(V)$, so dass die Punkte und P_0, P_1, P_2 paarweise verschieden sind. Dann liegen auch die Bildpunkte $\pi(P_0), \pi(P_1), \pi(P_2), \pi(P_3)$ auf einer projektiven Gerade in $P(V')$, die Punkte $\pi(P_0), \pi(P_1), \pi(P_2)$ sind paarweise verschieden und für das Doppelverhältnis gilt*

$$DV(\pi(P_0), \pi(P_1), \pi(P_2), \pi(P_3)) = DV(P_0, P_1, P_2, P_3).$$

BEWEIS. Es bezeichne G die eindeutige projektive Gerade in $P(V)$, die die Punkte P_0, P_1, P_2, P_3 enthält. Weiters sei $\tilde{\pi} : \mathbb{K}P^1 \rightarrow G$ die eindeutige projektive Abbildung mit

$$\tilde{\pi}([1 : 0]) = P_0, \quad \tilde{\pi}([0 : 1]) = P_1 \quad \text{und} \quad \tilde{\pi}([1 : 1]) = P_2$$

und $\tilde{\pi}([x_0 : x_1]) = P_3$, also $DV(P_0, P_1, P_2, P_3) = \frac{x_1}{x_0}$. Es ist dann $G' := \pi(G)$ die eindeutige projektive Gerade in $P(V')$, die die Punkte $\pi(P_0), \pi(P_1), \pi(P_2), \pi(P_3)$

enthält. Weiters ist $\tilde{\pi}' := \pi \circ \tilde{\pi}: \mathbb{K}P^1 \rightarrow G'$ die eindeutige projektive Abbildung, für die

$$\tilde{\pi}'([1 : 0]) = \pi(P_0), \quad \tilde{\pi}'([0 : 1]) = \pi(P_1) \quad \text{und} \quad \tilde{\pi}'([1 : 1]) = \pi(P_2)$$

gilt. Darüber hinaus haben wir $\tilde{\pi}'([x_0 : x_1]) = \pi(\tilde{\pi}([x_0 : x_1])) = \pi(P_3)$ und daher $DV(\pi(P_0), \pi(P_1), \pi(P_2), \pi(P_3)) = \frac{x_1}{x_0} = DV(P_0, P_1, P_2, P_3)$. \square

VIII.3.34. PROPOSITION. *Seien $P_0 = [x_0 : y_0]$, $P_1 = [x_1 : y_1]$, $P_2 = [x_2 : y_2]$, $P_3 = [x_3 : y_3]$ vier Punkte in $\mathbb{K}P^1$, sodass P_0, P_1, P_2 paarweise verschieden sind. Dann gilt für ihr Doppelverhältnis*

$$DV(P_0, P_1, P_2, P_3) = \frac{\begin{vmatrix} x_0 & x_3 \\ y_0 & y_3 \end{vmatrix}}{\begin{vmatrix} x_0 & x_2 \\ y_0 & y_2 \end{vmatrix}} : \frac{\begin{vmatrix} x_1 & x_3 \\ y_1 & y_3 \end{vmatrix}}{\begin{vmatrix} x_1 & x_2 \\ y_1 & y_2 \end{vmatrix}}$$

BEWEIS. Aus der Cramer'schen Regel erhalten wir

$$\begin{vmatrix} x_0 & x_1 \\ y_0 & y_1 \end{vmatrix} \begin{pmatrix} x_2 \\ y_2 \end{pmatrix} = \mu_0 \begin{pmatrix} x_0 \\ y_0 \end{pmatrix} + \mu_1 \begin{pmatrix} x_1 \\ y_1 \end{pmatrix},$$

wobei

$$\mu_0 = \begin{vmatrix} x_2 & x_1 \\ y_2 & y_1 \end{vmatrix} \quad \text{und} \quad \mu_1 = \begin{vmatrix} x_0 & x_2 \\ y_0 & y_2 \end{vmatrix}.$$

Die Matrix

$$A = \begin{pmatrix} \mu_0 x_0 & \mu_1 x_1 \\ \mu_0 y_0 & \mu_1 y_1 \end{pmatrix}$$

repräsentiert daher eine Projektivität $\pi: \mathbb{K}P^1 \rightarrow \mathbb{K}P^1$, sodass $\pi([1 : 0]) = P_0$, $\pi([0 : 1]) = P_1$ und $\pi([1 : 1]) = P_2$. Nochmaliges Anwenden der Cramer'schen Regel liefert

$$\det(A) \begin{pmatrix} x_3 \\ y_3 \end{pmatrix} = \lambda_0 \begin{pmatrix} \mu_0 x_0 \\ \mu_0 y_0 \end{pmatrix} + \lambda_1 \begin{pmatrix} \mu_1 x_1 \\ \mu_1 y_1 \end{pmatrix},$$

wobei

$$\lambda_0 = \begin{vmatrix} x_3 & \mu_1 x_1 \\ y_3 & \mu_1 y_1 \end{vmatrix} \quad \text{und} \quad \lambda_1 = \begin{vmatrix} \mu_0 x_0 & x_3 \\ \mu_0 y_0 & y_3 \end{vmatrix},$$

d.h. $\pi([\lambda_0 : \lambda_1]) = P_3$. Für das Doppelverhältnis erhalten wir

$$DV(P_0, P_1, P_2, P_3) = \frac{\lambda_1}{\lambda_0} = \frac{\mu_0 \begin{vmatrix} x_0 & x_3 \\ y_0 & y_3 \end{vmatrix}}{\mu_1 \begin{vmatrix} x_3 & x_1 \\ y_3 & y_1 \end{vmatrix}} = \frac{\begin{vmatrix} x_0 & x_3 \\ y_0 & y_3 \end{vmatrix}}{\begin{vmatrix} x_0 & x_2 \\ y_0 & y_2 \end{vmatrix}} : \frac{\begin{vmatrix} x_1 & x_3 \\ y_1 & y_3 \end{vmatrix}}{\begin{vmatrix} x_1 & x_2 \\ y_1 & y_2 \end{vmatrix}},$$

wie behauptet. \square

VIII.3.35. BEISPIEL. Sind $x_0, x_1, x_2, x_3 \in \mathbb{K}$ und bezeichnen

$$P_i = \iota(x_i) = [1 : x_i] \in \mathbb{K}P^1$$

die entsprechenden Punkte in $\mathbb{K}P^1$, dann gilt

$$DV(P_0, P_1, P_2, P_3) = \frac{x_0 - x_3}{x_1 - x_3} : \frac{x_0 - x_2}{x_1 - x_2}.$$

Dies folgt aus der vorangehenden Proposition.

VIII.3.36. BEMERKUNG (Teilverhältnis). Seien p_0 und p_1 zwei verschiedene Punkte auf einer affinen Geraden G . Es existiert daher ein eindeutiger affiner Isomorphismus $\alpha: \mathbb{K} \rightarrow G$, sodass $\alpha(0) = p_0$ und $\alpha(1) = p_1$, nämlich $\alpha(x) = p_0 + x(p_1 - p_0)$. Ist p_2 ein weiterer Punkt auf G , dann wird der eindeutig bestimmte Skalar $x \in \mathbb{K}$ mit $\alpha(x) = p_2$ als *Teilverhältnis*¹⁷ der Punkte p_0, p_1, p_2 bezeichnet,

$$\text{TV}(p_0, p_1, p_2) := \frac{p_2 - p_0}{p_1 - p_0}.$$

Beachte

$$\text{TV}(p_0, p_1, p_0) = 0 \quad \text{und} \quad \text{TV}(p_0, p_1, p_1) = 1,$$

sowie

$$\text{TV}(p_0, p_2, p_1) = \frac{1}{\text{TV}(p_0, p_1, p_2)} \quad \text{und} \quad \text{TV}(p_1, p_0, p_2) = 1 - \text{TV}(p_0, p_1, p_2).$$

Offensichtlich haben wir auch:

$$\text{TV}(p_0, p_1, p_3) = \text{TV}(p_0, p_1, p_2) \cdot \text{TV}(p_0, p_2, p_3). \quad (\text{VIII.16})$$

Für die Punkte 0, 1 und x auf der affinen Geraden \mathbb{K} gilt

$$\text{TV}(0, 1, x) = x.$$

Ist G' eine weitere affine Gerade und $\alpha: G \rightarrow G'$ ein affiner Isomorphismus, dann

$$\text{TV}(\alpha(p_0), \alpha(p_1), \alpha(p_2)) = \text{TV}(p_0, p_1, p_2),$$

d.h. affine Abbildungen bewahren das Teilverhältnis. Ist $P = G \cup \{\infty\}$ eine projektive Gerade und $p_i \in G$, dann gilt

$$\text{TV}(p_0, p_1, p_2) := \text{DV}(p_0, \infty, p_1, p_2)$$

und aus der vorangehenden Beispiel folgt

$$\text{DV}(p_0, p_1, p_2, p_3) = \frac{\text{TV}(p_3, p_1, p_0)}{\text{TV}(p_2, p_1, p_0)}.$$

Der letzten Gleichung verdankt das Doppelverhältnis seinen Namen.

VIII.3.37. PROPOSITION. Seien G_0, G_1, G_2, G_3 vier verschiedene projektive Geraden in einer projektiven Ebene E , die sich alle im Punkt P schneiden. Seien G und G' zwei weitere projektive Geraden in E , die P nicht enthalten. Weiters bezeichne P_i den eindeutigen Schnittpunkt von G mit G_i sowie P'_i den eindeutigen Schnittpunkt von G' mit G_i . Dann gilt

$$\text{DV}(P_0, P_1, P_2, P_3) = \text{DV}(P'_0, P'_1, P'_2, P'_3).$$

¹⁷In der Literatur finden sich auch andere Konventionen.

BEWEIS. Nach Proposition VIII.3.33 genügt es eine Projektivität $\pi: E \rightarrow E$ zu konstruieren, sodass $\pi(P_i) = P'_i$, $i = 0, \dots, 3$. Dafür wiederum genügt es eine Projektivität π zu konstruieren, sodass $\pi(G) = G'$ und $\pi(H) = H$, für jede projektive Gerade H durch P . Es bezeichne F die Gerade durch P und den Schnittpunkt von G und G' . Wir betrachten nun den affinen Raum $A := E \setminus F$. Nach Konstruktion sind die Geraden G und G' parallel in A , denn sie haben keinen Schnittpunkt in A . Nach Konstruktion sind auch alle Geraden durch P in A parallel. Eine geeignete Translation in Richtung dieser Geraden wird G auf G' abbilden und hat daher die gewünschte Eigenschaft. \square

Für zwei verschiedene Punkte p und q eines affinen Raums bezeichnen wir die von ihnen aufgespannte Gerade mit $\overline{pq} := \langle \{p, q\} \rangle_{\text{aff}}$. Der Strahlensatz, siehe Aufgabe 23, lässt sich wie folgt reformulieren:

VIII.3.38. PROPOSITION (Strahlensatz). *Seien g_1 und g_2 zwei verschiedene affine Geraden in einer affinen Ebene. Weiters seien p_1, q_1 zwei verschiedene Punkte auf g_1 und p_2, q_2 zwei verschiedenen Punkte auf g_2 .*

a) *Schneiden sich g_1 und g_2 in einem Punkt o und sind p_1, p_2, q_1, q_2 alle verschieden von o dann gilt:*

$$\overline{p_1 p_2} \parallel \overline{q_1 q_2} \quad \Leftrightarrow \quad \text{TV}(o, p_1, q_1) = \text{TV}(o, p_2, q_2).$$

b) *Haben g_1 und g_2 keinen Schnittpunkt dann gilt:*

$$\overline{p_1 p_2} \parallel \overline{q_1 q_2} \quad \Leftrightarrow \quad q_1 - p_1 = q_2 - p_2.$$

Daraus erhalten wir sofort:

VIII.3.39. PROPOSITION (Satz von Desargues, affine Version). *Seien g_1, g_2, g_3 drei verschiedene affine Geraden in einer affinen Ebene, die sich entweder in einem Punkt o schneiden oder alle parallel sind. Weiters seien p_i und q_i verschiedene Punkte auf g_i , die im Fall $g_1 \cap g_2 \cap g_3 = \{o\}$ auch alle verschieden von o sein sollen, $i = 1, 2, 3$. Dann gilt:*

$$\overline{p_1 p_2} \parallel \overline{q_1 q_2} \quad \text{und} \quad \overline{p_2 p_3} \parallel \overline{q_2 q_3} \quad \Rightarrow \quad \overline{p_1 p_3} \parallel \overline{q_1 q_3}.$$

Für zwei verschiedene Punkte P und Q eines projektiven Raums bezeichnen wir die von ihnen aufgespannte Gerade ebenfalls mit $\overline{PQ} := \langle \{P, Q\} \rangle_{\text{proj}}$.

VIII.3.40. SATZ (Satz von Desargues). *Seien G_1, G_2, G_3 drei verschiedene projektive Geraden in einer projektiven Ebene E , die sich in einem Punkt o schneiden. Weiters seien P_i und Q_i verschiedene Punkte auf G_i , die auch alle verschieden von o sein sollen, $i = 1, 2, 3$. Dann liegen die Schnittpunkte*

$$R_3 := \overline{P_1 P_2} \cap \overline{Q_1 Q_2}, \quad R_1 := \overline{P_2 P_3} \cap \overline{Q_2 Q_3} \quad \text{und} \quad R_2 := \overline{P_1 P_3} \cap \overline{Q_1 Q_3}$$

auf einer Geraden.

BEWEIS. Es bezeichne F die eindeutige Gerade durch R_3 und R_1 . Wir betrachten nun den affinen Raum $A := E \setminus F$. Nach Konstruktion sind die Geraden $\overline{P_1P_2}$ und $\overline{Q_1Q_2}$ in A parallel. Ebenso sind die Geraden $\overline{P_2P_3}$ und $\overline{Q_2Q_3}$ in A parallel. Nach Proposition VIII.3.39 sind dann auch die Geraden $\overline{P_1P_3}$ und $\overline{Q_1Q_3}$ parallel in A , ihr Schnittpunkt R_2 liegt daher ebenfalls in F . Dies zeigt, dass die Punkte R_1, R_2, R_3 alle auf der Geraden F liegen. \square

VIII.3.41. PROPOSITION (Satz von Pappos, affine Version). *Seien g und g' zwei verschiedene affine Geraden in einer affinen Ebene. Weiters seien p_1, p_2, p_3 verschiedene Punkte auf g und p'_1, p'_2, p'_3 verschiedene Punkte auf g' . Schneiden sich g und g' in einem Punkt o , dann seien die Punkte p_i und p'_i auch alle verschieden von o , $i = 1, 2, 3$. In dieser Situation gilt:*

$$\overline{p_1p'_3} \parallel \overline{p_3p'_1} \quad \text{und} \quad \overline{p_1p'_2} \parallel \overline{p_2p'_1} \quad \Rightarrow \quad \overline{p_2p'_3} \parallel \overline{p_3p'_2}.$$

BEWEIS. Wir betrachten zunächst den Fall wo sich g und g' in einem Punkt o schneiden. Aus dem Strahlensatz erhalten wir

$$\text{TV}(o, p_1, p_3) = \text{TV}(o, p'_3, p'_1) \quad \text{und} \quad \text{TV}(o, p_1, p_2) = \text{TV}(o, p'_2, p'_1).$$

Mit (VIII.16) folgt:

$$\text{TV}(o, p_2, p_3) = \frac{\text{TV}(o, p_1, p_3)}{\text{TV}(o, p_1, p_2)} = \frac{\text{TV}(o, p'_3, p'_1)}{\text{TV}(o, p'_2, p'_1)} = \text{TV}(o, p'_3, p'_2).$$

Nach dem Strahlensatz sind die Geraden $\overline{p_2p'_3}$ und $\overline{p_3p'_2}$ daher parallel. Haben die beiden Geraden g und g' keinen Schnittpunkt, dann gilt

$$p_3 - p_1 = p'_1 - p'_3 \quad \text{und} \quad p_2 - p_1 = p'_1 - p'_2$$

und daher $p_3 - p_2 = p'_2 - p'_3$. Wieder folgt, dass $\overline{p_2p'_3}$ und $\overline{p_3p'_2}$ parallel sind. \square

VIII.3.42. SATZ (Satz von Pappos). *Seien G und G' zwei verschiedene projektive Geraden in einer projektiven Ebene E , die sich im Punkt o schneiden. Weiters seien P_1, P_2, P_3 verschiedene Punkte auf G und P'_1, P'_2, P'_3 verschiedene Punkte auf G' , die auch alle verschieden von o seien sollen, $i = 1, 2, 3$. Dann liegen die Schnittpunkte*

$$R_2 := \overline{P_1P'_3} \cap \overline{P_3P'_1}, \quad R_3 := \overline{P_1P'_2} \cap \overline{P_2P'_1} \quad \text{und} \quad R_1 := \overline{P_2P'_3} \cap \overline{P_3P'_2}$$

auf einer Geraden.

BEWEIS. Es bezeichne F die eindeutige Gerade durch R_2 und R_3 . Wir betrachten wieder den affinen Raum $A := E \setminus F$. Nach Konstruktion sind die Geraden $\overline{P_1P'_3}$ und $\overline{P_3P'_1}$ parallel in A , und auch die Geraden $\overline{P_1P'_2}$ und $\overline{P_2P'_1}$ sind parallel in A . Nach Proposition VIII.3.41 sind daher auch $\overline{P_2P'_3}$ und $\overline{P_3P'_2}$ parallel in A , ihr Schnittpunkt R_1 liegt also in F . Dies zeigt, dass die drei Punkte R_1, R_2, R_3 alle auf der Geraden F liegen. \square

VIII.3.43. DEFINITION (Dualität). Sei V ein endlich dimensionaler Vektorraum. Ist $P = P(W)$ ein projektiver Teilraum in $P(V)$, dann wird $P^\circ := P(W^\circ)$ der zu P duale projektive Teilraum von $P(V^*)$ genannt. Hier bezeichnet $W^\circ \subseteq V^*$ den Annihilator des linearen Teilraums $W \subseteq V$, d.h. $W^\circ = \{\alpha \in V^* : \alpha|_W = 0\}$.

VIII.3.44. LEMMA. Seien P und Q projektive Teilräume von $P(V)$. Dann gilt:

- (a) $\dim(P) + \dim(P^\circ) + 1 = \dim(P(V)) = \dim(P(V^*))$.
- (b) $P \subseteq Q \Leftrightarrow P^\circ \supseteq Q^\circ$.
- (c) $(P \cap Q)^\circ = \langle P^\circ \cup Q^\circ \rangle_{proj}$.
- (d) $\langle P \cup Q \rangle_{proj} = P^\circ \cap Q^\circ$.

BEWEIS. Für jeden linearen Teilraum W von V gilt die Dimensionsformel $\dim(W^\circ) = \dim(V) - \dim(W)$ und somit (a). Für je zwei lineare Teilräume W und U von V gilt $W \subseteq U \Leftrightarrow W^\circ \supseteq U^\circ$, also (b). Nach Satz III.4.9 gilt weiters $(W \cap U)^\circ = W^\circ + U^\circ$ und $(W + U)^\circ = W^\circ \cap U^\circ$. Daraus folgen (c) und (d). \square

Daraus erhalten wir sofort folgendes metamathematische Dualitätsprinzip:

VIII.3.45. SATZ (Dualitätsprinzip). Angenommen wir haben einen Satz über projektive Teilräume eines n -dimensionalen projektiven Raums bewiesen, dessen Formulierung nur die Begriffe Dimension, Inklusion, Durchschnitt und projektive Hülle involviert. Ersetzen wir in dieser Aussage überall

- (a) k -dimensionaler Teilraum durch $(n - k - 1)$ -dimensionaler Teilraum,
 - (b) enthält durch ist enthalten in,
 - (c) Durchschnitt von Teilräumen durch projektive Hülle ihrer Vereinigung und
 - (d) projektive Hülle der Vereinigung durch Durchschnitt der Teilräume,
- so erhalten wir erneut einen Satz über projektive Teilräume eines n -dimensionalen projektiven Raums.

VIII.3.46. BEISPIEL. Die Aussage "Zwei verschiedene Punkte einer projektiven Ebene liegen auf genau einer projektiven Geraden" ist dual zu der Aussage "Zwei verschiedene projektive Geraden einer projektiven Ebene haben genau einen Schnittpunkt."

VIII.3.47. BEISPIEL. Die Aussage "Zwei verschiedene Punkte eines 3-dimensionalen projektiven Raums liegen auf genau einer projektiven Geraden" ist dual zu der Aussage "Zwei verschiedene projektive Ebenen eines 3-dimensionalen projektiven Raums schneiden sich in einer projektiven Geraden."

VIII.3.48. BEISPIEL. Die Aussage "Zwei verschiedene Punkte eines n -dimensionalen projektiven Raums liegen auf genau einer projektiven Geraden" ist dual zu der Aussage "Zwei verschiedene projektive Hyperebenen eines projektiven Raums schneiden sich in einem $(n - 2)$ -dimensionalen projektiven Teilraum."

Dual zum Satz von Desargues ist folgende Umkehrung desselben:

VIII.3.49. KOROLLAR. In einer projektiven Ebene betrachte drei verschiedene Punkte R_1, R_2 und R_3 , die auf einer projektiven Geraden F liegen. Weiters seien H_i und K_i verschiedene projektive Geraden durch R_i , die auch alle verschieden von F sein sollen, $i = 1, 2, 3$. Betrachten wir die Schnittpunkte

$$\begin{aligned} P_3 &:= H_1 \cap H_2, & P_2 &:= H_1 \cap H_3, & P_1 &:= H_2 \cap H_3, \\ Q_3 &:= K_1 \cap K_2, & Q_2 &:= K_1 \cap K_3, & Q_1 &:= K_2 \cap K_3, \end{aligned}$$

dann schneiden sich die Geraden $\overline{P_1Q_1}$, $\overline{P_2Q_2}$ und $\overline{P_3Q_3}$ in einem Punkt.

Dualisieren wir den Satz von Pappus, so erhalten wir:

VIII.3.50. KOROLLAR. Seien R und R' zwei verschiedene Punkte einer projektiven Ebene. Weiters seien G_1, G_2, G_3 drei Geraden durch R und G'_1, G'_2, G'_3 drei Geraden durch R' . Betrachten wir die Schnittpunkte

$$\begin{aligned} P_1 &:= G_2 \cap G'_3, & P_2 &:= G_1 \cap G'_3, & P_3 &:= G_1 \cap G'_2 \\ Q_1 &:= G'_2 \cap G_3, & Q_2 &:= G'_1 \cap G_3, & Q_3 &:= G'_1 \cap G_2 \end{aligned}$$

dann schneiden sich die Geraden $\overline{P_1Q_1}$, $\overline{P_2Q_2}$ und $\overline{P_3Q_3}$ in einem Punkt.

VIII.4. Projektive Quadriken. Sei V ein endlich dimensionaler Vektorraum über einem Körper \mathbb{K} mit $\text{char}(\mathbb{K}) \neq 2$.

VIII.4.1. DEFINITION (Projektive Quadriken). Unter einer projektiven Quadrik verstehen wir jede Teilmenge $E \subseteq P(V)$ von der Form

$$E = \{[v] \in P(V) : q(v) = 0\},$$

wobei q eine quadratische Form auf V bezeichnet. Beachte, dass q keine Funktion auf $P(V)$ definiert, die Nullstellenmenge aber trotzdem wohldefiniert ist, denn aus $q(v) = 0$ folgt $q(\lambda v) = 0$, für jeden Skalar $\lambda \in \mathbb{K}$.

VIII.4.2. BEISPIEL. Die Teilmenge

$$E := \{[x_0 : x_1 : x_2] \in \mathbb{RP}^2 \mid x_0^2 + x_1^2 - x_2^2 = 0\}$$

bildet eine Quadrik in \mathbb{RP}^2 . Bezeichnet $\iota: \mathbb{R}^2 \rightarrow \mathbb{RP}^2$, $\iota\left(\begin{smallmatrix} x_1 \\ x_2 \end{smallmatrix}\right) = [1 : x_1 : x_2]$, die Standardeinbettung, so ist der affine Teil, $\iota^{-1}(E) = \left\{\begin{smallmatrix} x_1 \\ x_2 \end{smallmatrix}\right\} \in \mathbb{R}^2 \mid 1 + x_1^2 - x_2^2 = 0\}$, eine Hyperbel in \mathbb{R}^2 . Beachte, dass E aber auch die beiden Fernpunkte $[0 : 1 : 1]$ und $[0 : 1 : -1]$ enthält, sie entsprechen den Richtungen der beiden Asymptoten der Hyperbel.

VIII.4.3. PROPOSITION. Ist $\pi: P(V) \rightarrow P(V')$ eine projektive Abbildung und E eine projektive Quadrik in $P(V)$, dann ist auch $\pi(E)$ eine projektive Quadrik in $P(V')$.

BEWEIS. Nach Voraussetzung existiert ein linearer Isomorphismus $\phi: V \rightarrow V'$, sodass $\pi([v]) = [\phi(v)]$ für alle $0 \neq v \in V$. Weiters existiert eine quadratische Form q auf V , sodass $E = \{[v] \in P(V) : q(v) = 0\}$. Es ist daher $q' := q \circ \phi^{-1}$ eine quadratische Form auf V' , und somit $E' := \{[v'] \in P(V') : q'(v') = 0\}$ eine projektive Quadrik in $P(V')$. Nach Konstruktion gilt $\pi(E) = E'$. \square

VIII.4.4. PROPOSITION. Sei $A \subseteq V$ eine affine Hyperebene, $0 \notin A$ und $\iota: A \rightarrow P(V)$ die damit assoziierte Einbettung. Ist E eine projektive Quadrik in $P(V)$, dann ist $\iota^{-1}(E)$ eine affine Quadrik in A . Umgekehrt ist jede affine Quadrik in A von der Form $\iota^{-1}(E)$ für eine geeignete, i.A. nicht eindeutig bestimmte, projektive Quadrik E in $P(V)$.

BEWEIS. Ist $E \subseteq P(V)$ eine projektive Quadrik, dann existiert eine quadratische Form q auf V , sodass $E = \{[v] \in P(V) : q(v) = 0\}$. Die Einschränkung $Q := q|_A: A \rightarrow \mathbb{K}$ ist eine quadratische Funktion. Offensichtlich gilt $\iota^{-1}(E) = \{a \in A : Q(a) = 0\}$, also ist $\iota^{-1}(E)$ eine affine Quadrik in A .

Für die zweite Aussage nehmen wir o.B.d.A. $P(V) = \mathbb{K}P^n$ und $\iota: \mathbb{K}^n \rightarrow \mathbb{K}P^n$, $\iota(x) = [1 : x_1 : \dots : x_n]$, an. Ist $Q(x) = x^t Ax + b^t x + c$ eine quadratische Funktion auf \mathbb{K}^n , dann repräsentiert die symmetrische $(n+1) \times (n+1)$ -Matrix

$$\begin{pmatrix} c & b^t/2 \\ b/2 & A \end{pmatrix}$$

eine quadratische Form $q: \mathbb{K}^{n+1} \rightarrow \mathbb{K}$, und es gilt

$$\iota^{-1}(\{[v] \in \mathbb{K}P^n : q(v) = 0\}) = \{x \in \mathbb{K}^n : Q(x) = 0\},$$

denn $\begin{pmatrix} 1 \\ x \end{pmatrix}^t \begin{pmatrix} c & b^t/2 \\ b/2 & A \end{pmatrix} \begin{pmatrix} 1 \\ x \end{pmatrix} = x^t Ax + b^t x + c = Q(x)$. □

VIII.4.5. BEISPIEL. Betrachte die affine Quadrik

$$E_0 := \left\{ \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \in \mathbb{R}^2 : x_2 = x_1^2 + 1 \right\}.$$

Wir wollen eine projektive Quadrik $E \subseteq \mathbb{R}P^2$ bestimmen, deren affiner Teil mit E_0 überein stimmt, d.h. so, dass $\iota^{-1}(E) = E_0$ gilt, wobei $\iota: \mathbb{R}^2 \rightarrow \mathbb{R}P^2$ die Standardeinbettung bezeichnet. Ersetzen wir in der definierenden Gleichung für E_0 die Variable x_i durch $\frac{x_i}{x_0}$ so erhalten wir

$$\frac{x_2}{x_0} = \left(\frac{x_1}{x_0} \right)^2 + 1.$$

Multiplikation mit x_0^2 liefert den Ausdruck

$$x_2 x_0 = x_1^2 + x_0^2.$$

Die gesuchte projektive Quadrik ist daher

$$E = \{[x_0 : x_1 : x_2] \in \mathbb{R}P^2 \mid x_2 x_0 = x_1^2 + x_0^2\}.$$

VIII.4.6. SATZ (Normalform komplexer projektiver Quadriken). Sei V ein endlich dimensionaler komplexer Vektorraum, $\dim(P(V)) = n$ und E eine projektive Quadrik in $P(V)$. Dann existiert eine projektive Abbildung $\pi: \mathbb{C}P^n \rightarrow P(V)$, die E auf die Quadrik

$$\pi^{-1}(E) = \{[z_0 : \dots : z_n] \in \mathbb{C}P^n \mid z_0^2 + \dots + z_{k-1}^2 = 0\}$$

abbildet, für ein geeignetes $0 \leq k \leq n+1$.

BEWEIS. Nach Voraussetzung existiert eine quadratische Form q auf V , sodass $E = \{[v] \in P(V) : q(v) = 0\}$. Nach Satz VII.1.25 existiert ein linearer Isomorphismus $\phi: \mathbb{C}^{n+1} \rightarrow V$, sodass $q(\phi(z)) = z_0^2 + \dots + z_{k-1}^2$, wobei $k = \text{rank}(q)$. Die davon repräsentierte projektive Abbildung $\pi = P_\phi: \mathbb{C}P^n \rightarrow P(V)$ hat die gewünschte Eigenschaft. \square

VIII.4.7. SATZ (Normalform reeller projektiver Quadriken). *Sei V ein endlich dimensionaler reeller Vektorraum, $\dim(P(V)) = n$ und E eine projektive Quadrik in $P(V)$. Dann existiert eine projektive Abbildung $\pi: \mathbb{R}P^n \rightarrow P(V)$, die E auf die Quadrik*

$$\pi^{-1}(E) = \{[x_0 : \dots : x_n] \in \mathbb{R}P^n \mid x_0^2 + \dots + x_{p-1}^2 - x_p^2 - \dots - x_{p+k-1}^2 = 0\}$$

abbildet, für geeignete $0 \leq k \leq p \leq n+1$.

BEWEIS. Nach Voraussetzung existiert eine quadratische Form q auf V , sodass $E = \{[v] \in P(V) : q(v) = 0\}$. Nach Satz VII.1.35 existiert ein linearer Isomorphismus $\phi: \mathbb{R}^{n+1} \rightarrow V$, sodass

$$q(\phi(x)) = x_0^2 + \dots + x_{p-1}^2 - x_p^2 - \dots - x_{p+k-1}^2,$$

wobei (p, k) die Signatur von q bezeichnet. Da q und $-q$ die selbe projektive Quadrik beschreiben, dürfen wir o.B.d.A. $p \geq k$ annehmen. Die davon repräsentierte projektive Abbildung, $\pi = P_\phi: \mathbb{R}P^n \rightarrow P(V)$, hat die gewünschte Eigenschaft. \square

VIII.4.8. BEISPIEL. Jede Quadrik in der projektiven Ebene $\mathbb{R}P^2$ ist zu einer der folgenden Quadriken projektiv kongruent:

- (a) $\{[x_0 : x_1 : x_2] \in \mathbb{R}P^2 \mid 0 = 0\}$ (ganz $\mathbb{R}P^2$)
- (b) $\{[x_0 : x_1 : x_2] \in \mathbb{R}P^2 \mid x_0^2 = 0\}$ (Gerade)
- (c) $\{[x_0 : x_1 : x_2] \in \mathbb{R}P^2 \mid x_0^2 + x_1^2 = 0\}$ (ein Punkt)
- (d) $\{[x_0 : x_1 : x_2] \in \mathbb{R}P^2 \mid x_0^2 - x_1^2 = 0\}$ (zwei Geraden)
- (e) $\{[x_0 : x_1 : x_2] \in \mathbb{R}P^2 \mid x_0^2 + x_1^2 + x_2^2 = 0\}$ (leer)
- (f) $E := \{[x_0 : x_1 : x_2] \in \mathbb{R}P^2 \mid x_0^2 + x_1^2 - x_2^2 = 0\}$

Bezeichnet $\iota_0: \mathbb{R}^2 \rightarrow \mathbb{R}P^2$, $\iota_0(x, y) = [1 : x : y]$, die Standardeinbettung, dann ist $\iota_0^{-1}(E) = \{(\frac{x}{y}) \in \mathbb{R}^2 : 1 + x^2 - y^2 = 0\}$ eine Hyperbel und es liegen zwei Punkte von E auf der Ferngeraden, $[0 : 1 : 1]$ und $[0 : 1 : -1]$. Betrachten wir die Einbettung $\iota_2: \mathbb{R}^2 \rightarrow \mathbb{R}P^2$, $\iota_2(x, y) = [x : y : 1]$, dann ist $\iota_2^{-1}(E) = \{(\frac{x}{y}) \in \mathbb{R}^2 : x^2 + y^2 - 1 = 0\}$ ein Kreis und kein Punkt von E liegt auf der Ferngeraden. Für die mit der affinen Hyperebene $\left(\begin{smallmatrix} 0 \\ 1 \\ -1 \end{smallmatrix}\right) + \left\langle \begin{smallmatrix} 1 \\ 0 \\ 0 \end{smallmatrix}\right\rangle, \left(\begin{smallmatrix} 0 \\ 1 \\ 1 \end{smallmatrix}\right)\right\rangle$ assoziierten Einbettung $\iota: \mathbb{R}^2 \rightarrow \mathbb{R}P^2$, $\iota(x, y) = [x : y + 1 : y - 1]$, ist $\iota^{-1}(E) = \{(\frac{x}{y}) \in \mathbb{R}^2 : x^2 + 4y = 0\}$ eine Parabel. In diesem Fall liegt nur ein Punkt von E auf der Ferngeraden, nämlich $[0 : 1 : 1]$.

VIII.4.9. BEISPIEL. Jede Quadrik in des projektiven Raums $\mathbb{R}P^3$ ist zu einer der folgenden Quadriken projektiv kongruent:

- (a) $\{[x_0 : x_1 : x_2 : x_3] \in \mathbb{R}P^3 \mid 0 = 0\}$ (ganz $\mathbb{R}P^3$)

- (b) $\{[x_0 : x_1 : x_2 : x_3] \in \mathbb{RP}^3 \mid x_0^2 = 0\}$ (eine Ebene)
 (c) $\{[x_0 : x_1 : x_2 : x_3] \in \mathbb{RP}^3 \mid x_0^2 + x_1^2 = 0\}$ (eine Gerade)
 (d) $\{[x_0 : x_1 : x_2 : x_3] \in \mathbb{RP}^3 \mid x_0^2 - x_1^2 = 0\}$ (zwei Ebenen)
 (e) $\{[x_0 : x_1 : x_2 : x_3] \in \mathbb{RP}^3 \mid x_0^2 + x_1^2 + x_2^2 = 0\}$ (ein Punkt)
 (f) $\{[x_0 : x_1 : x_2 : x_3] \in \mathbb{RP}^3 \mid x_0^2 + x_1^2 - x_2^2 = 0\}$
 (g) $\{[x_0 : x_1 : x_2 : x_3] \in \mathbb{RP}^3 \mid x_0^2 + x_1^2 + x_2^2 + x_3^2 = 0\}$ (leer)
 (h) $\{[x_0 : x_1 : x_2 : x_3] \in \mathbb{RP}^3 \mid x_0^2 + x_1^2 + x_2^2 - x_3^2 = 0\}$
 (i) $\{[x_0 : x_1 : x_2 : x_3] \in \mathbb{RP}^3 \mid x_0^2 + x_1^2 - x_2^2 - x_3^2 = 0\}$

VIII.4.10. BEISPIEL. Wir wollen eine Projektivität $\pi: \mathbb{RP}^2 \rightarrow \mathbb{RP}^2$ bestimmen, die die projektive Quadrik

$$E = \{[x : y : z] \in \mathbb{RP}^2 \mid x^2 + 2y^2 + z^3 + 2xy + 2xz + 4yz = 0\}$$

auf Normalform bringt. Durch Ergänzen auf vollständige Quadrate erhalten wir

$$x^2 + 2y^2 + z^3 + 2xy + 2xz + 4yz = (x + y + z)^2 + (y + z)^2 - z^2.$$

Die lineare Abbildung $\phi: \mathbb{R}^3 \rightarrow \mathbb{R}^3$, $\phi \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x+y+z \\ y+z \\ z \end{pmatrix}$, induziert daher eine Projektivität $\pi = P_\phi: \mathbb{RP}^2 \rightarrow \mathbb{RP}^2$, sodass

$$\pi(E) = \{[x : y : z] \in \mathbb{RP}^2 \mid x^2 + y^2 - z^2 = 0\}$$

Normalform hat.

IX. Multilineare Algebra

Bevor wir uns im zweiten Abschnitt mit dem Tensorprodukt auseinandersetzen, wollen wir zum Aufwärmen zwei Konstruktionen mit Vektorräumen besprechen, die durch universelle Eigenschaften charakterisiert werden können: das Produkt und die direkte Summe von Vektorräumen.

Wir orientieren uns an der Darstellung in [11, Chapter 4].

IX.1. Produkte und direkte Summen. Alle Vektorräume in diesem Kapitel seien über einem fixen aber beliebigen Körper \mathbb{K} definiert. Unter dem *Produkt* zweier Vektorräume V und W verstehen wir die Menge aller Paare $V \times W$ mit der komponentenweisen Addition und Skalarmultiplikation:

$$(v, w) + (v', w') := (v + v', w + w'), \quad \lambda(v, w) := (\lambda v, \lambda w),$$

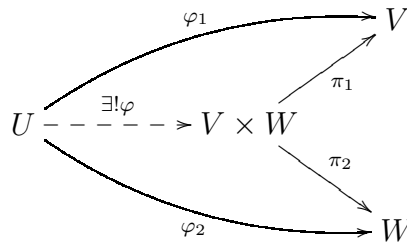
wobei $v, v' \in V$, $w, w' \in W$ und $\lambda \in \mathbb{K}$. Es lässt sich leicht verifizieren, dass dadurch $V \times W$ zu einem Vektorraum über \mathbb{K} wird. Sind V und W beide endlich dimensional, dann gilt

$$\dim(V \times W) = \dim V + \dim W.$$

Die linearen Abbildungen,

$$\pi_1: V \times W \rightarrow V \quad \text{und} \quad \pi_2: V \times W \rightarrow W,$$

$\pi_1(v, w) := v$, $\pi_2(v, w) := w$ werden als *kanonische Projektionen* bezeichnet, und besitzen folgende universelle Eigenschaft: *Ist U ein weiterer Vektorraum und sind $\varphi_1: U \rightarrow V$ und $\varphi_2: U \rightarrow W$ zwei lineare Abbildungen, dann existiert eine eindeutige lineare Abbildung $\varphi: U \rightarrow V \times W$, sodass $\pi_1 \circ \varphi = \varphi_1$ und $\pi_2 \circ \varphi = \varphi_2$. Offensichtlich leistet nämlich $\varphi(u) := (\varphi_1(u), \varphi_2(u))$ das Gewünschte. Diese *universelle Eigenschaft des Produkts* lässt sich übersichtlich durch folgendes kommutative Diagramm veranschaulichen:*



Die beiden Projektionen liefern lineare Abbildungen

$$L(U, V \times W) \xrightarrow{(\pi_1)_*} L(U, V) \quad \text{und} \quad L(U, V \times W) \xrightarrow{(\pi_2)_*} L(U, W),$$

$(\pi_1)_*(\varphi) := \pi_1 \circ \varphi$ und $(\pi_2)_*(\varphi) := \pi_2 \circ \varphi$. Diese induzieren eine lineare Abbildung $L(U, V \times W) \rightarrow L(U, V) \times L(U, W)$. Aus der universellen Eigenschaft des Produkts folgt sofort, dass dies ein linearer Isomorphismus ist. Wir erhalten somit einen kanonischen Isomorphismus

$$L(U, V \times W) = L(U, V) \times L(U, W),$$

wobei die Projektionen des Produkts auf der rechten Seite den Abbildungen $(\pi_1)_*$ und $(\pi_2)_*$ entsprechen.

Analog lassen sich auch Produkte endlich vieler Vektorräume V_1, \dots, V_n konstruieren. Die Menge $V_1 \times \dots \times V_n$ bildet bezüglich komponentenweiser Addition und Skalarmultiplikation einen Vektorraum, und die kanonischen (linearen) Projektionen

$$\pi_i: V_1 \times \dots \times V_n \rightarrow V_i, \quad \pi_i(v_1, \dots, v_n) := v_i,$$

haben folgende universelle Eigenschaft: Ist U ein Vektorraum und sind $\varphi_i: U \rightarrow V_i$ lineare Abbildungen, $i = 1, \dots, n$, dann existiert genau eine lineare Abbildung $\varphi: U \rightarrow V_1 \times \dots \times V_n$, sodass $\pi_i \circ \varphi = \varphi_i$, für jedes $i = 1, \dots, n$. Offensichtlich leistet nämlich $\varphi(u) := (\varphi_1(u), \dots, \varphi_n(u))$ das Gewünschte. Wieder induzieren die linearen Abbildungen $(\pi_i)_*: L(U, \prod_{j=1}^n V_j) \rightarrow L(U, V_i)$ einen kanonischen linearen Isomorphismus

$$L\left(U, \prod_{i=1}^n V_i\right) = \prod_{i=1}^n L(U, V_i),$$

sodass die Projektionen der rechten Seite den Abbildungen $(\pi_i)_*$ entsprechen.

Allgemeiner sei I eine beliebige Indexmenge und V_i ein Vektorraum, für jedes $i \in I$. Das Produkt

$$\prod_{i \in I} V_i = \{v = (v_i)_{i \in I} \mid \forall i \in I : v_i \in V_i\},$$

ist selbst ein Vektorraum bezüglich komponentenweise Addition und Skalarmultiplikation, $(v+w)_i := v_i + w_i$, $(\lambda v)_i := \lambda v_i$. Für $i \in I$ wird die lineare Abbildung $\pi_i: \prod_{j \in I} V_j \rightarrow V_i$, $\pi_i(v) := v_i$, als *kanonische Projektion* auf den i -ten Faktor bezeichnet. Diese haben folgende Eigenschaft: *Ist U ein Vektorraum und sind $\varphi_i: U \rightarrow V_i$ linear für jedes $i \in I$, dann existiert eine eindeutige lineare Abbildung $\varphi: U \rightarrow \prod_{i \in I} V_i$, sodass $\pi_i \circ \varphi = \varphi_i$, für jedes $i \in I$. Die Abbildung $\varphi(u) := (\varphi_i(u))_{i \in I}$ hat nämlich die gewünschte Eigenschaft. Das Produkt $\prod_{i \in I} V_i$ zusammen mit den kanonischen Projektionen π_i ist durch diese *universelle Eigenschaft des Produkts* im wesentlichen eindeutig bestimmt. Genauer haben wir:*

IX.1.1. LEMMA. *Seien $p_i: P \rightarrow V_i$ lineare Abbildungen, $i \in I$, mit der universellen Eigenschaft des Produkts, d.h. sind $\varphi_i: U \rightarrow V_i$ linear für jedes $i \in I$, dann existiere eine eindeutige lineare Abbildung $\varphi: U \rightarrow P$, sodass $p_i \circ \varphi = \varphi_i$, für jedes $i \in I$. Dann gibt es einen eindeutig bestimmten linearen Isomorphismus $\phi: P \xrightarrow{\cong} \prod_{i \in I} V_i$, sodass $\pi_i \circ \phi = p_i$, für jedes $i \in I$.*

BEWEIS. Aus der universellen Eigenschaft des Produkts erhalten wir eine eindeutig bestimmte lineare Abbildung $\phi: P \rightarrow \prod_{i \in I} V_i$, sodass $\pi_i \circ \phi = p_i$, für jedes $i \in I$. Analog erhalten wir aus der universellen Eigenschaft der Abbildungen p_i eine lineare Abbildung $\psi: \prod_{i \in I} V_i \rightarrow P$, sodass $p_i \circ \psi = \pi_i$, für jedes $i \in I$. Es folgt $p_i \circ \psi \circ \phi = p_i = p_i \circ \text{id}_P$, für alle $i \in I$, und daher $\psi \circ \phi = \text{id}_P$ wegen der Eindeutigkeitsaussage in der universellen Eigenschaft von P . Analog gilt $\pi_i \circ \phi \circ \psi = \pi_i = \pi_i \circ \text{id}_{\prod_{i \in I} V_i}$, für alle $i \in I$, und daher $\phi \circ \psi = \text{id}_{\prod_{i \in I} V_i}$ wegen der Eindeutigkeitsaussage in der universellen Eigenschaft von $\prod_{i \in I} V_i$. Folglich ist ϕ ein Isomorphismus.

$\psi = \pi_i = \pi_i \circ \text{id}_{\prod_{i \in I} V_i}$, also $\phi \circ \psi = \text{id}_{\prod_{i \in I} V_i}$ aufgrund der Eindeutigkeitsaussage in der universellen Eigenschaft des Produkts. Dies zeigt, dass ϕ und ψ zueinander inverse Isomorphismen sind. \square

Die Projektionen $\pi_i: \prod_{j \in I} V_j \rightarrow V_i$ liefern lineare Abbildungen

$$(\pi_i)_*: L\left(U, \prod_{j \in I} V_j\right) \rightarrow L(U, V_i),$$

$(\pi_i)_*(\varphi) = \pi_i \circ \varphi$, und diese induzieren einen kanonischen linearen Isomorphismus

$$L\left(U, \prod_{i \in I} V_i\right) = \prod_{i \in I} L(U, V_i),$$

sodass die Projektionen der rechten Seiten den Abbildungen $(\pi_i)_*$ entsprechen.

Sind $\varphi_i: V_i \rightarrow W_i$ linear für jedes $i \in I$, dann erhalten wir aus der universellen Eigenschaft des Produkts eine eindeutig bestimmte lineare Abbildung $\varphi: \prod_{i \in I} V_i \rightarrow \prod_{i \in I} W_i$, sodass $\pi_i^W \circ \varphi = \varphi_i \circ \pi_i^V$, für jedes $i \in I$. Dabei bezeichnen $\pi_i^V: \prod_{j \in I} V_j \rightarrow V_i$ und $\pi_i^W: \prod_{j \in I} W_j \rightarrow W_i$ die kanonischen Projektionen der beiden Produkte. Diese Abbildung φ wird üblicherweise mit $\prod_{i \in I} \varphi_i$ bezeichnet,

$$\prod_{i \in I} \varphi_i: \prod_{i \in I} V_i \rightarrow \prod_{i \in I} W_i.$$

Aus der Eindeutigkeit erhalten wir sofort

$$\prod_{i \in I} \psi_i \circ \prod_{i \in I} \varphi_i = \prod_{i \in I} (\psi_i \circ \varphi_i) \quad \text{und} \quad \prod_{i \in I} \text{id}_{V_i} = \text{id}_{\prod_{i \in I} V_i},$$

wobei $\psi_i: W_i \rightarrow U_i$ weitere lineare Abbildungen bezeichnen. Im Fall endlicher Produkte, schreiben wir dafür:

$$V_1 \times \cdots \times V_n \xrightarrow{\varphi_1 \times \cdots \times \varphi_n} W_1 \times \cdots \times W_n.$$

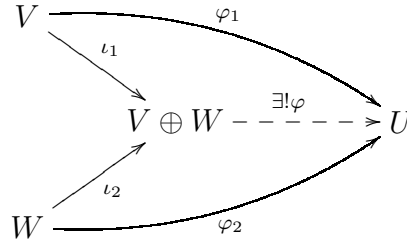
Offensichtlich gilt $(\varphi_1 \times \cdots \times \varphi_n)(v_1, \dots, v_n) = (\varphi_1(v_1), \dots, \varphi_n(v_n))$, für $v_i \in V_i$.

Auch die (äußere) *direkte Summe* $V \oplus W$ zweier Vektorräume V und W hat als zugrundeliegenden Vektorraum die Menge der Paare (v, w) mit der komponentenweisen Addition und Skalarmultiplikation. Wir betrachten nun aber die beiden *kanonischen Inklusionen*

$$\iota_1: V \rightarrow V \oplus W \quad \text{und} \quad \iota_2: W \rightarrow V \oplus W,$$

$\iota_1(v) := (v, 0)$ und $\iota_2(w) := (0, w)$. Diese haben folgende Eigenschaft: *Ist U ein weiterer Vektorraum und sind $\varphi_1: V \rightarrow U$ und $\varphi_2: W \rightarrow U$ zwei lineare Abbildungen, dann existiert genau eine lineare Abbildung $\varphi: V \oplus W \rightarrow U$, sodass $\varphi \circ \iota_1 = \varphi_1$ und $\varphi \circ \iota_2 = \varphi_2$. Offensichtlich leistet nämlich $\varphi(v, w) = \varphi_1(v) +$*

$\varphi_2(w)$ das Gewünschte. Diagrammatisch lässt sich diese *universelle Eigenschaft der direkten Summe* wie folgt veranschaulichen:



Die beiden Inklusionen liefern lineare Abbildungen

$$L(V \oplus W, U) \xrightarrow{\iota_1^*} L(V, U) \quad \text{und} \quad L(V \oplus W, U) \xrightarrow{\iota_2^*} L(W, U),$$

$\iota_1^*(\varphi) = \varphi \circ \iota_1$, $\iota_2^*(\varphi) = \varphi \circ \iota_2$. Diese induzieren einen kanonischen linearen Isomorphismus

$$L(V \oplus W, U) = L(V, U) \times L(W, U),$$

sodass die Projektionen der rechten Seite den Abbildungen ι_1^* und ι_2^* entsprechen.

IX.1.2. BEMERKUNG. Sind V und W komplementäre Teilräume von U , dann induzieren die Inklusionen einen Isomorphismus zwischen der äußeren direkten Summe $V \oplus W$ und U , vgl. Proposition II.5.5(c).

Analog definieren wir die äußere direkte Summe $V_1 \oplus \cdots \oplus V_n$ endlich vieler Vektorräume V_1, \dots, V_n als die Menge aller Tupel (v_1, \dots, v_n) mit der komponentenweisen Addition und Skalarmultiplikation, und betrachten die kanonischen Inklusionen

$$\iota_i: V_i \rightarrow V_1 \oplus \cdots \oplus V_n, \quad \iota_i(v_i) = (0, \dots, v_i, 0, \dots, 0),$$

wobei der nicht-triviale Eintrag an der i -ten Stelle stehen soll. Diese haben folgende Eigenschaft: *Sind $\varphi_i: V_i \rightarrow U$ lineare Abbildungen, $i = 1, \dots, n$, dann existiert eine eindeutig bestimmte lineare Abbildung $\varphi: V_1 \oplus \cdots \oplus V_n \rightarrow U$, sodass $\varphi \circ \iota_i = \varphi_i$, für jedes $i = 1, \dots, n$.*

Allgemeiner, seien V_i Vektorräume, $i \in I$, wobei I eine beliebige Indexmenge bezeichnet. Dann ist die Menge aller endlichen Tupel,

$$\bigoplus_{i \in I} V_i := \left\{ v \in \prod_{i \in I} V_i : \text{nur endlich viele } v_i \neq 0 \right\},$$

ein Teilraum des Produkts und daher selbst ein Vektorraum bezüglich komponentenweiser Addition und Skalarmultiplikation. Für $i \in I$ wird die lineare Abbildung

$$\iota_i: V_i \rightarrow \bigoplus_{j \in I} V_j, \quad (\iota_i(v))_j := \begin{cases} v & \text{falls } i = j \\ 0 & \text{andernfalls} \end{cases}$$

als kanonische Inklusion des i -ten Summanden bezeichnet. Diese haben folgende universelle Eigenschaft: *Sind $\varphi_i: V_i \rightarrow U$ lineare Abbildungen, $i \in I$, dann existiert genau eine lineare Abbildung $\varphi: \bigoplus_{i \in I} V_i \rightarrow U$, sodass $\varphi \circ \iota_i = \varphi_i$, für alle*

$i \in I$. Die lineare Abbildung $\varphi(v) = \sum_{i \in I} \varphi_i(v_i)$ leistet nämlich das Gewünschte. Beachte, dass dies eine endliche Summe ist, da nur endlich viele v_i verschieden von 0 sind. Die direkte Summe $\bigoplus_{i \in I}$ zusammen mit den kanonischen Inklusionen ι_i ist durch die *universelle Eigenschaft der direkten Summe* im wesentlichen eindeutig bestimmt. Genauer gilt:

IX.1.3. LEMMA. *Seien $j_i: V_i \rightarrow S$ lineare Abbildungen mit der universellen Eigenschaft der direkten Summe, d.h. sind $\varphi_i: V_i \rightarrow U$ lineare Abbildungen, so existiere genau eine lineare Abbildung $\varphi: S \rightarrow U$, sodass $\varphi \circ j_i = \varphi_i$, für alle $i \in I$. Dann gibt es genau einen linearen Isomorphismus $\phi: \bigoplus_{i \in I} V_i \xrightarrow{\cong} S$, sodass $\phi \circ \iota_i = j_i$, für jedes $i \in I$.*

Der Beweis ist völlig analog zu dem von Lemma IX.1.1, vgl. Aufgabe 52.

Die Inklusionen liefern lineare Abbildungen $\iota_i^*: L(\bigoplus_{j \in I} V_j, U) \rightarrow L(V_i, U)$, $\iota_i^*(\varphi) = \varphi \circ \iota_i$, und diese induzieren einen kanonischen linearen Isomorphismus

$$L\left(\bigoplus_{i \in I} V_i, U\right) = \prod_{i \in I} L(V_i, U), \quad (\text{IX.1})$$

sodass die Projektionen der rechten Seite den Abbildungen ι_i^* entsprechen. Für $U = \mathbb{K}$ liefert dies einen kanonischen Isomorphismus

$$\left(\bigoplus_{i \in I} V_i\right)^* = \prod_{i \in I} V_i^*.$$

Sind $\varphi_i: V_i \rightarrow W_i$ linear für jedes $i \in I$, dann erhalten wir aus der universellen Eigenschaft der direkten Summe eine eindeutig bestimmte lineare Abbildung $\varphi: \bigoplus_{i \in I} V_i \rightarrow \bigoplus_{i \in I} W_i$, sodass $\varphi \circ \iota_i^V = \iota_i^W \circ \varphi_i$, für jedes $i \in I$. Dabei bezeichnen $\iota_i^V: V_i \rightarrow \bigoplus_{j \in I} V_j$ und $\iota_i^W: W_i \rightarrow \bigoplus_{j \in I} W_j$ die kanonischen Inklusionen der beiden Summen. Diese Abbildung φ wird üblicherweise mit $\bigoplus_{i \in I} \varphi_i$ bezeichnet,

$$\bigoplus_{i \in I} \varphi_i: \bigoplus_{i \in I} V_i \rightarrow \bigoplus_{i \in I} W_i.$$

Aus der Eindeutigkeit erhalten wir sofort

$$\bigoplus_{i \in I} \psi_i \circ \bigoplus_{i \in I} \varphi_i = \bigoplus_{i \in I} (\psi_i \circ \varphi_i) \quad \text{und} \quad \bigoplus_{i \in I} \text{id}_{V_i} = \text{id}_{\bigoplus_{i \in I} V_i},$$

wobei $\psi_i: W_i \rightarrow U_i$ weitere lineare Abbildungen bezeichnen. Im Fall endlicher Produkte, schreiben wir dafür:

$$V_1 \oplus \cdots \oplus V_n \xrightarrow{\varphi_1 \oplus \cdots \oplus \varphi_n} W_1 \oplus \cdots \oplus W_n.$$

Offensichtlich gilt $(\varphi_1 \oplus \cdots \oplus \varphi_n)(v_1, \dots, v_n) = (\varphi_1(v_1), \dots, \varphi_n(v_n))$, für $v_i \in V_i$.

IX.2. Tensorprodukt. Eine Abbildung $\mu: V_1 \times \cdots \times V_k \rightarrow W$ heißt multilinear oder k -linear, falls $\mu(v_1, \dots, v_k)$ linear in jeder Eintragung ist. Die Menge dieser Abbildungen bildet einen Vektorraum, den wir mit $L_k(V_1, \dots, V_k; W)$ bezeichnen. Beachte,

$$L_{k+r}(V_1, \dots, V_{k+r}; W) = L_k(V_1, \dots, V_k; L_r(V_{k+1}, \dots, V_{k+r}; W)), \quad (\text{IX.2})$$

via $\mu(v_1, \dots, v_k)(v_{k+1}, \dots, v_{k+r}) = \mu(v_1, \dots, v_{k+r})$.

IX.2.1. BEISPIEL. $L_2(V, V; \mathbb{K})$ ist der Vektorraum der Bilinearformen auf V . Die Determinante kann als $\det \in L_n(\mathbb{K}^n, \dots, \mathbb{K}^n; \mathbb{K})$ aufgefasst werden.

IX.2.2. SATZ. Sind V und W zwei Vektorräume über einem Körper \mathbb{K} , dann existiert ein \mathbb{K} -Vektorraum T und eine bilineare Abbildung $t: V \times W \rightarrow T$ mit folgender Eigenschaft: Für jeden Vektorraum U und jede bilineare Abbildung $b: V \times W \rightarrow U$ existiert eine eindeutige lineare Abbildung $\varphi: T \rightarrow U$, sodass $b = \varphi \circ t$. Der Vektorraum T zusammen mit der bilinearen Abbildung t sind durch diese Eigenschaft bis auf kanonischen Isomorphismus eindeutig bestimmt, d.h. ist $\tilde{t}: V \times W \rightarrow \tilde{T}$ eine weitere bilineare Abbildung mit obiger Eigenschaft, dann existiert ein eindeutiger linearer Isomorphismus $\phi: T \xrightarrow{\cong} \tilde{T}$, sodass $\tilde{t} = \phi \circ t$.

BEWEIS. Wir beginnen mit der Eindeutigkeit. Seien also $t: V \times W \rightarrow T$ und $\tilde{t}: V \times W \rightarrow \tilde{T}$ zwei bilineare Abbildungen, die beide die im Satz formulierte Eigenschaft besitzen. Es existiert daher eine eindeutige lineare Abbildung $\phi: T \rightarrow \tilde{T}$, sodass $\phi \circ t = \tilde{t}$. Analog existiert eine lineare Abbildung $\psi: \tilde{T} \rightarrow T$, sodass $\psi \circ \tilde{t} = t$. Es folgt $\psi \circ \phi \circ t = t = \text{id}_T \circ t$. Aus der Eindeutigkeitsaussage in der universellen Eigenschaft von T erhalten wir daraus $\psi \circ \phi = \text{id}_T$. Völlig analog lässt sich $\phi \circ \psi = \text{id}_{\tilde{T}}$ zeigen, denn $\psi \circ \phi \circ \tilde{t} = \tilde{t} = \text{id}_{\tilde{T}} \circ \tilde{t}$. Somit sind ϕ und ψ zueinander inverse Isomorphismen.

Nun zur Existenz einer bilinearen Abbildung $t: V \times W \rightarrow T$ mit der gewünschten Eigenschaft. Wir betrachten den Vektorraum, $F(V \times W, \mathbb{K}) = \mathbb{K}^{V \times W}$ aller Funktionen $V \times W \rightarrow \mathbb{K}$ mit der üblichen punktweisen Addition und Skalarmultiplikation. Für $v \in V$ und $w \in W$ bezeichne $\delta_{v,w} \in F(V \times W, \mathbb{K})$ die Funktion mit $\delta_{v,w}(v, w) = 1$ und $\delta_{v,w}(\tilde{v}, \tilde{w}) = 0$ falls $\tilde{v} \neq v$ oder $\tilde{w} \neq w$. Es bezeichne $M \subseteq F(V \times W, \mathbb{K})$ den von den Elementen

$$\delta_{v,w}, \quad v \in V, w \in W$$

aufgespannten Teilraum. Weiters bezeichne $R \subseteq M$ den von den Elementen

$$\begin{aligned} \delta_{v_1+v_2,w} - \delta_{v_1,w} - \delta_{v_2,w}, & \quad \delta_{\lambda v,w} - \lambda \delta_{v,w} \\ \delta_{v,w_1+w_2} - \delta_{v,w_1} - \delta_{v,w_2}, & \quad \delta_{v,\lambda w} - \lambda \delta_{v,w} \end{aligned}$$

aufgespannten Teilraum, wobei $v, v_1, v_2 \in V$, $w, w_1, w_2 \in W$ und $\lambda \in \mathbb{K}$. Schließlich sei $T := M/R$ und $t: V \times W \rightarrow T$, $t(v, w) := [\delta_{v,w}]$. Da im Quotientenraum

M/R die Relationen

$$\begin{aligned} [\delta_{v_1+v_2,w}] &= [\delta_{v_1,w}] + [\delta_{v_2,w}], & [\delta_{\lambda v,w}] &= \lambda[\delta_{v,w}] \\ [\delta_{v,w_1+w_2}] &= [\delta_{v,w_1}] + [\delta_{v,w_2}], & [\delta_{v,\lambda w}] &= \lambda[\delta_{v,w}] \end{aligned}$$

gelten, ist t bilinear. Es bleibt zu zeigen, dass t die universelle Eigenschaft besitzt. Sei also $b: V \times W \rightarrow U$ bilinear. Da die Funktionen $\delta_{v,w}$ eine Basis von M bilden, existiert eine eindeutige lineare Abbildung $\tilde{\varphi}: M \rightarrow U$, sodass $\tilde{\varphi}(\delta_{v,w}) = b(v, w)$, für alle $v \in V$ und $w \in W$. Aus der Bilinearität von b folgt $\tilde{\varphi}|_R = 0$, etwa ist

$$\begin{aligned} \tilde{\varphi}(\delta_{v_1+v_2,w} - \delta_{v_1,w} - \delta_{v_2,w}) &= \tilde{\varphi}(\delta_{v_1+v_2,w}) - \tilde{\varphi}(\delta_{v_1,w}) - \tilde{\varphi}(\delta_{v_2,w}) \\ &= b(v_1 + v_2, w) - b(v_1, w) - b(v_2, w) = 0, \end{aligned}$$

und analog lässt sich zeigen, dass $\tilde{\varphi}$ auf den anderen Erzeugern von R verschwindet. Somit existiert eine lineare Abbildung $\varphi: T \rightarrow U$, sodass $\varphi([f]) = \tilde{\varphi}(f)$, für jedes $f \in M$. Insbesondere erhalten wir $\varphi(t(v, w)) = \varphi([\delta_{v,w}]) = \tilde{\varphi}(\delta_{v,w}) = b(v, w)$, und somit $\varphi \circ t = b$. Nach Konstruktion bildet $\{t(v, w) : v \in V, w \in W\}$ ein Erzeugendensystem von T . Daraus folgt sofort, dass die lineare Abbildung $\varphi: T \rightarrow U$ durch die Gleichung $\varphi \circ t = b$ eindeutig bestimmt ist. \square

IX.2.3. DEFINITION (Tensorprodukt). Seien V und W zwei Vektorräume über einem Körper \mathbb{K} . Unter einem (dem) *Tensorprodukt* von V und W verstehen wir einen \mathbb{K} -Vektorraum $V \otimes W$ zusammen mit einer bilinearen Abbildung

$$V \times W \xrightarrow{\otimes} V \otimes W, \quad (v, w) \mapsto v \otimes w,$$

die folgende Eigenschaft hat: *Für jeden \mathbb{K} -Vektorraum U und jede bilineare Abbildung $b: V \times W \rightarrow U$ existiert eine eindeutige lineare Abbildung $\varphi: V \otimes W \rightarrow U$, sodass $\varphi(v \otimes w) = b(v, w)$, für alle $v \in V$ und $w \in W$:*

$$\begin{array}{ccc} V \times W & \xrightarrow{\otimes} & V \otimes W \\ & \searrow b & \swarrow \exists! \varphi \\ & & U \end{array}$$

Dies wird als *universelle Eigenschaft des Tensorprodukts* bezeichnet. Nach dem vorangehenden Satz existiert das Tensorprodukt beliebiger Vektorräume und ist im wesentlichen eindeutig.¹⁸

Aufgrund der Bilinearität der Abbildung $V \times W \xrightarrow{\otimes} V \otimes W$ gilt:

- (a) $(v_1 + v_2) \otimes w = v_1 \otimes w + v_2 \otimes w$
- (b) $v \otimes (w_1 + w_2) = v \otimes w_1 + v \otimes w_2$
- (c) $\lambda(v \otimes w) = (\lambda v) \otimes w = v \otimes (\lambda w)$

¹⁸Wir könnten expliziter auch $V \otimes W := T = M/R$ definieren, vgl. Satz IX.2.2.

für alle $v, v_i \in V$, $w, w_i \in W$ und $\lambda \in \mathbb{K}$. Die universelle Eigenschaft des Tensorprodukts liefert eine Bijektion

$$L_2(V, W; U) = L(V \otimes W, U).$$

Aus der Eindeutigkeitsaussage in der universellen Eigenschaft folgt sofort, dass dies ein linearer Isomorphismus ist. Da auch $L_2(V, W; U) = L(V, L(W, U))$, siehe (IX.2), erhalten wir einen kanonischen linearen Isomorphismus

$$L(V \otimes W, U) = L(V, L(W, U)). \quad (\text{IX.3})$$

IX.2.4. BEISPIEL. Ist V ein Vektorraum über \mathbb{K} , dann existiert ein kanonischer Isomorphismus

$$V \otimes \mathbb{K} = V, \quad v \otimes \lambda \leftrightarrow \lambda v.$$

Um dies einzusehen, fassen wir die Skalarmultiplikation als bilineare Abbildung $V \times \mathbb{K} \rightarrow V$ auf, und erhalten eine lineare Abbildung $\phi: V \otimes \mathbb{K} \rightarrow V$, sodass $\phi(v \otimes \lambda) = \lambda v$, für alle $v \in V$ und $\lambda \in \mathbb{K}$. Andererseits ist auch die Abbildung $\psi: V \rightarrow V \otimes \mathbb{K}$, $\psi(v) := v \otimes 1$, linear. Es folgt $\phi(\psi(v)) = \phi(v \otimes 1) = v$, also $\phi \circ \psi = \text{id}_V$ aber auch $\psi(\phi(v \otimes \lambda)) = \psi(\lambda v) = (\lambda v) \otimes 1 = v \otimes \lambda$, für alle $v \in V$ und $\lambda \in \mathbb{K}$, also $\psi \circ \phi = \text{id}_{V \otimes \mathbb{K}}$ wegen der Eindeutigkeitsaussage in der universellen Eigenschaft des Tensorprodukts. Somit sind ϕ und ψ zueinander inverse Isomorphismen.

IX.2.5. PROPOSITION. Seien V und W zwei Vektorräume über \mathbb{K} , B eine Basis von V und C eine Basis von W . Dann bildet $\{b \otimes c : b \in B, c \in C\}$ eine Basis von $V \otimes W$. Sind V und W beide endlich dimensional, dann ist auch $V \otimes W$ endlich dimensional, und es gilt

$$\dim(V \otimes W) = \dim(V) \cdot \dim(W).$$

BEWEIS. Es bezeichne $L \subseteq V \otimes W$ den von $\{b \otimes c : b \in B, c \in C\}$ aufgespannten Teilraum, und $\pi: V \otimes W \rightarrow (V \otimes W)/L$ die kanonische Projektion. Es gilt daher $\pi(b \otimes c) = 0$ für alle $b \in B$ und $c \in C$. Da B und C Erzeugendensysteme von V bzw. W sind, folgt $\pi(v \otimes w) = 0$, für alle $v \in V$ und $w \in W$. Sind nämlich $v = \sum_{b \in B} \lambda_b b$ und $w = \sum_{c \in C} \mu_c c$ Darstellungen, wobei fast alle λ_b und μ_c verschwinden, dann folgt $\pi(v \otimes w) = \sum_{b \in B, c \in C} \lambda_b \mu_c \pi(b \otimes c) = 0$. Aus der Eindeutigkeitsaussage in der universellen Eigenschaft des Tensorprodukts erhalten wir $\pi = 0$, also $L = V \otimes W$. Dies zeigt, dass $\{b \otimes c : b \in B, c \in C\}$ ein Erzeugendensystem von $V \otimes W$ bildet.

Um auch die lineare Unabhängigkeit zu zeigen, sei

$$\sum_{b \in B, c \in C} \lambda_{b,c} b \otimes c = 0, \quad (\text{IX.4})$$

wobei nur endlich viele der Skalare $\lambda_{b,c} \in \mathbb{K}$ verschieden von Null sind. Seien $b_0 \in B$ und $c_0 \in C$ fix. Da B eine Basis ist, existiert ein lineares Funktional $\beta: V \rightarrow \mathbb{K}$, sodass $\beta(b_0) = 1$ und $\beta(b) = 0$, für alle $b \in B \setminus \{b_0\}$. Genauso existiert ein lineares Funktional $\gamma: W \rightarrow \mathbb{K}$, sodass $\gamma(c_0) = 1$ und $\gamma(c) = 0$, für alle $c \in C \setminus \{c_0\}$.

Da $V \times W \rightarrow \mathbb{K}$, $(v, w) \mapsto \beta(v)\gamma(w)$, bilinear ist, existiert nach der universellen Eigenschaft des Tensorprodukts ein lineares Funktional $\varphi: V \otimes W \rightarrow \mathbb{K}$, sodass $\varphi(v \otimes w) = \beta(v)\gamma(w)$, für alle $v \in V$ und $w \in W$. Insbesondere gilt für $b \in B$ und $c \in C$:

$$\varphi(b \otimes c) = \begin{cases} 1 & \text{falls } b = b_0 \text{ und } c = c_0, \\ 0 & \text{andernfalls.} \end{cases}$$

Wenden wir dieses Funktional φ auf (IX.4) an, erhalten wir $\lambda_{b_0, c_0} = 0$. Dies zeigt, dass $\{b \otimes c : b \in B, c \in C\}$ auch linear unabhängig in $V \otimes W$ ist. \square

IX.2.6. BEMERKUNG. Elemente der Form $v \otimes w$ in $V \otimes W$ werden als *elementare Tensoren* bezeichnet. Nach dem vorangehenden Resultat, lässt sich jedes Element in $V \otimes W$ als endliche Summe elementarer Tensoren schreiben. I.A. ist aber nicht jedes Element des Tensorprodukts von der Form $v \otimes w$, vgl. Aufgabe 56.

Sind $\varphi: V \rightarrow V'$ und $\psi: W \rightarrow W'$ zwei lineare Abbildungen, dann existiert eine eindeutige lineare Abbildung $\varphi \otimes \psi: V \otimes W \rightarrow V' \otimes W'$, sodass

$$(\varphi \otimes \psi)(v \otimes w) = \varphi(v) \otimes \psi(w), \quad (\text{IX.5})$$

für alle $v \in V$ und $w \in W$. Dies folgt aus der universellen Eigenschaft des Tensorprodukts, denn der Ausdruck $\varphi(v) \otimes \psi(w)$ ist bilinear in v und w .

IX.2.7. PROPOSITION. Für $\varphi, \varphi_1, \varphi_2 \in L(V, V')$, $\psi, \psi_1, \psi_2 \in L(W, W')$, $\varphi' \in L(V', V'')$, $\psi' \in L(W', W'')$ und $\lambda \in \mathbb{K}$ gilt:

- (a) $\text{id}_V \otimes \text{id}_W = \text{id}_{V \otimes W}$
- (b) $(\varphi' \circ \varphi) \otimes (\psi' \circ \psi) = (\varphi' \otimes \psi') \circ (\varphi \otimes \psi)$
- (c) $(\varphi_1 + \varphi_2) \otimes \psi = \varphi_1 \otimes \psi + \varphi_2 \otimes \psi$
- (d) $\varphi \otimes (\psi_1 + \psi_2) = \varphi \otimes \psi_1 + \varphi \otimes \psi_2$
- (e) $(\lambda\varphi) \otimes \psi = \lambda(\varphi \otimes \psi) = \varphi \otimes (\lambda\psi)$

BEWEIS. Dies folgt leicht daraus, dass die Abbildung $\varphi \otimes \psi$ durch (IX.5) eindeutig bestimmt ist, vgl. Aufgabe 55. Etwa gilt für alle $v \in V$ und $w \in W$,

$$\begin{aligned} (\varphi_1 \otimes \psi + \varphi_2 \otimes \psi)(v \otimes w) &= (\varphi_1 \otimes \psi)(v \otimes w) + (\varphi_2 \otimes \psi)(v \otimes w) \\ &= \varphi_1(v) \otimes \psi(w) + \varphi_2(v) \otimes \psi(w) \\ &= (\varphi_1(v) + \varphi_2(v)) \otimes \psi(w) \\ &= (\varphi_1 + \varphi_2)(v) \otimes \psi(w) \\ &= ((\varphi_1 + \varphi_2) \otimes \psi)(v \otimes w), \end{aligned}$$

und wegen der eben angesprochenen Eindeutigkeit daher (c). \square

Das Tensorprodukt ist in folgendem Sinn distributiv:

IX.2.8. PROPOSITION. Sei I eine beliebige Indexmenge, V_i Vektorräume, $i \in I$, und bezeichnen $\iota_i: V_i \rightarrow \bigoplus_{j \in I} V_j$ die kanonischen Inklusionen. Dann induzieren die linearen Abbildungen $V_i \otimes W \xrightarrow{\iota_i \otimes \text{id}_W} \left(\bigoplus_{j \in I} V_j \right) \otimes W$ einen kanonischen

linearen Isomorphismus

$$\bigoplus_{i \in I} (V_i \otimes W) = \left(\bigoplus_{i \in I} V_i \right) \otimes W.$$

Insbesondere gilt: $(V_1 \oplus V_2) \otimes W = (V_1 \otimes W) \oplus (V_2 \otimes W)$. Analog

$$V \otimes \left(\bigoplus_{i \in I} W_i \right) = \bigoplus_{i \in I} (V \otimes W_i),$$

und speziell: $V \otimes (W_1 \oplus W_2) = (V \otimes W_1) \oplus (V \otimes W_2)$.

BEWEIS. Aus (IX.1) und (IX.3) erhalten wir:

$$\begin{aligned} L\left(\left(\bigoplus_{i \in I} V_i\right) \otimes W, U\right) &= L\left(\bigoplus_{i \in I} V_i, L(W, U)\right) \\ &= \prod_{i \in I} L(V_i, L(W, U)) = \prod_{i \in I} L(V_i \otimes W, U). \end{aligned}$$

Sind $\varphi_i: V_i \otimes W \rightarrow U$ linear, $i \in I$, dann existiert also genau eine lineare Abbildung $\varphi: \left(\bigoplus_{i \in I} V_i\right) \otimes W \rightarrow U$, sodass $\varphi \circ (\iota_i \otimes \text{id}_W) = \varphi_i$, für alle $i \in I$. Die Abbildungen $\iota_i \otimes \text{id}_W: V_i \otimes W \rightarrow \left(\bigoplus_{j \in I} V_j\right) \otimes W$ haben also die universelle Eigenschaft der direkten Summe $V_i \otimes W \rightarrow \bigoplus_{j \in I} (V_j \otimes W)$, und induzieren daher einen Isomorphismus:

$$\bigoplus_{i \in I} (V_i \otimes W) = \left(\bigoplus_{i \in I} V_i \right) \otimes W.$$

Die zweite Aussage lässt sich völlig analog behandeln. \square

IX.2.9. PROPOSITION. Sei $0 \rightarrow V_1 \xrightarrow{\iota} V_2 \xrightarrow{\pi} V_3 \rightarrow 0$ eine kurze exakte Sequenz von Vektorräumen, und W ein weiterer Vektorraum. Dann ist auch

$$0 \rightarrow V_1 \otimes W \xrightarrow{\iota \otimes \text{id}_W} V_2 \otimes W \xrightarrow{\pi \otimes \text{id}_W} V_3 \otimes W \rightarrow 0$$

eine kurze exakte Sequenz.¹⁹ Für jeden Teilraum U von V kann daher $U \otimes W$ in kanonischer Weise als Teilraum von $V \otimes W$ aufgefasst werden, und wir haben einen kanonischen Isomorphismus:

$$(V \otimes W)/(U \otimes W) = (V/U) \otimes W.$$

BEWEIS. Da π surjektiv ist, existiert eine lineare Abbildung $\sigma: V_3 \rightarrow V_2$, sodass $\pi \circ \sigma = \text{id}_{V_3}$. Da, $\pi \circ (\text{id}_{V_2} - \sigma \circ \pi) = \pi - \pi \circ \sigma \circ \pi = \pi - \pi = 0$, hat die Abbildung $\text{id}_{V_2} - \sigma \circ \pi$ Werte in $\ker(\pi) = \text{img}(\iota)$, es existiert daher eine lineare Abbildung $\rho: V_2 \rightarrow V_1$, sodass $\iota \circ \rho = \text{id}_{V_2} - \sigma \circ \pi$. Es folgt $\iota \circ \rho \circ \iota = \iota - \sigma \circ \pi \circ \iota = \iota - 0 = \iota \circ \text{id}_{V_1}$, also $\rho \circ \iota = \text{id}_{V_1}$ denn ι ist injektiv. Zusammenfassend gilt:

$$\rho \circ \iota = \text{id}_{V_1}, \quad \iota \circ \rho + \sigma \circ \pi = \text{id}_{V_2} \quad \text{und} \quad \pi \circ \sigma = \text{id}_{V_3}.$$

¹⁹Eine Sequenz linearer Abbildungen $0 \rightarrow V_1 \xrightarrow{\iota} V_2 \xrightarrow{\pi} V_3 \rightarrow 0$ heißt exakt, falls ι injektiv ist, π surjektiv ist und $\text{img}(\iota) = \ker(\pi)$ gilt.

Mit den Rechenregeln in Proposition IX.2.7 folgt:

$$\begin{aligned}(\rho \otimes \text{id}_W) \circ (\iota \otimes \text{id}_W) &= \text{id}_{V_1 \otimes W}, & (\pi \otimes \text{id}_W) \circ (\sigma \otimes \text{id}_W) &= \text{id}_{V_3 \otimes W} \\ (\iota \otimes \text{id}_W) \circ (\rho \otimes \text{id}_W) + (\sigma \otimes \text{id}_W) \circ (\pi \otimes \text{id}_W) &= \text{id}_{V_2 \otimes W}\end{aligned}$$

Aus der ersten Gleichung folgt, dass $\iota \otimes \text{id}_W: V_1 \otimes W \rightarrow V_2 \otimes W$ injektiv ist. Aus der zweiten sehen wir, dass $\pi \otimes \text{id}_W: V_2 \otimes W \rightarrow V_3 \otimes W$ surjektiv ist. Aus der dritten folgt, $\ker(\pi \otimes \text{id}_W) \subseteq \text{img}(\iota \otimes \text{id}_W)$. Schließlich gilt auch die umgekehrte Inklusion, $\text{img}(\iota \otimes \text{id}_W) \subseteq \ker(\pi \otimes \text{id}_W)$, denn $(\pi \otimes \text{id}_W) \circ (\iota \otimes \text{id}_W) = (\pi \circ \iota) \otimes \text{id}_W = 0 \otimes \text{id}_W = 0$, da ja $\pi \circ \iota = 0$. Die verbleibenden Aussagen erhalten wir, indem wir das eben Bewiesene auf die kurze exakte Sequenz $0 \rightarrow U \rightarrow V \rightarrow V/U \rightarrow 0$ anwenden. \square

IX.2.10. PROPOSITION. *Seien V und W zwei endlich dimensionale Vektorräume. Dann induziert die bilineare Abbildung $V^* \times W \rightarrow L(V, W)$, $(\alpha, w) \mapsto \alpha w$, einen kanonischen linearen Isomorphismus:*

$$V^* \otimes W = L(V, W).$$

Darüber hinaus existiert ein kanonischer linearer Isomorphismus

$$(V \otimes W)^* = V^* \otimes W^*, \quad (\text{IX.6})$$

sodass $(\alpha \otimes \beta)(v \otimes w) = \alpha(v)\beta(w)$, für alle $\alpha \in V^*$, $\beta \in W^*$, $v \in V$ und $w \in W$.

BEWEIS. Sei b_1, \dots, b_n eine Basis von V und b_1^*, \dots, b_n^* die dazu duale Basis von V^* , d.h. $b_i^*(b_j) = \delta_{i,j}$. Weiters sei c_1, \dots, c_m eine Basis von W . Nach Proposition IX.2.5 ist daher $b_i^* \otimes c_j$, $1 \leq i \leq n$, $1 \leq j \leq m$, eine Basis von $V^* \otimes W$. Andererseits lässt sich jede lineare Abbildung $\varphi: V \rightarrow W$ in der Form $\varphi = \sum_{i=1}^n \sum_{j=1}^m A_{j,i} b_i^* c_j$ schreiben, wobei die eindeutig bestimmten Koeffizienten $A_{j,i} \in \mathbb{K}$ gerade die Eintragungen der Matrix von φ bezüglich der Basen b_i und c_j sind. Folglich ist $b_i^* c_j$, $1 \leq i \leq n$, $1 \leq j \leq m$, eine Basis von $L(V, W)$. Die Abbildung $V^* \times W \rightarrow L(V, W)$ bildet also eine Basis von $V^* \otimes W$ auf eine Basis von $L(V, W)$ ab, und ist daher ein Isomorphismus. Mit (IX.3) erhalten wir daraus

$$(V \otimes W)^* = L(V \otimes W, \mathbb{K}) = L(V, L(W, \mathbb{K})) = L(V, W^*) = V^* \otimes W^*,$$

und somit auch die zweite Behauptung. \square

IX.2.11. BEISPIEL. Für jeden endlich dimensionalen Vektorraum V erhalten wir aus dem vorangehenden Resultat einen kanonischen linearen Isomorphismus:

$$\text{end}(V) = V^* \otimes V.$$

Die identische Abbildung $\text{id}_V \in \text{end}(V)$ entspricht dabei dem Element

$$\sum_{i=1}^n b_i^* \otimes b_i \in V^* \otimes V,$$

wobei b_1, \dots, b_n eine beliebige Basis von V , und b_1^*, \dots, b_n^* die dazu duale Basis von V^* bezeichnen, d.h. $b_i^*(b_j) = \delta_{i,j}$. Die Spur, $\text{tr}: \text{end}(V) \rightarrow \mathbb{K}$, liefert ein lineares Funktional $\text{tr}: V^* \otimes V \rightarrow \mathbb{K}$, das wiederum der kanonischen bilinearen Abbildung $V^* \times V \rightarrow \mathbb{K}$, $(\alpha, v) \mapsto \alpha(v)$ entspricht.

IX.2.12. BEISPIEL. Sei X eine Menge und bezeichne $F_c(X)$ die Menge aller Funktionen $f: X \rightarrow \mathbb{K}$ mit endlichem Träger, d.h. $\{x \in X : f(x) \neq 0\}$ ist endlich. Sei Y eine weitere Menge und bezeichnen

$$p_1: X \times Y \rightarrow X, \quad p_2: X \times Y \rightarrow Y$$

die beiden Projektionen, $p_1(x, y) = x$, $p_2(x, y) = y$. Dann ist

$$F_c(X) \times F_c(Y) \rightarrow F_c(X \times Y), \quad (f, g) \mapsto (f \circ p_1) \cdot (g \circ p_2),$$

offensichtlich bilinear und induziert daher eine lineare Abbildung

$$F_c(X) \otimes F_c(Y) \rightarrow F_c(X \times Y), \quad (f \otimes g)(x, y) = f(x)g(y), \quad (\text{IX.7})$$

wobei $f \in F_c(X)$, $g \in F_c(Y)$, $x \in X$ und $y \in Y$. Beachte, dass $\{\delta_x : x \in X\}$ eine Basis von $F_c(X)$ und $\{\delta_y : y \in Y\}$ eine Basis von $F_c(Y)$ ist. Nach Proposition IX.2.5 bildet $\{\delta_x \otimes \delta_y : x \in X, y \in Y\}$ also eine Basis von $F_c(X) \otimes F_c(Y)$. Weiters ist $\{\delta_{(x,y)} : (x, y) \in X \times Y\}$ eine Basis von $F_c(X \times Y)$. Da $\delta_{(x,y)} = \delta_x \otimes \delta_y$ bildet (IX.7) also eine Basis auf eine Basis ab, und liefert daher einen kanonischen linearen Isomorphismus:

$$F_c(X) \otimes F_c(Y) = F_c(X \times Y).$$

Analog lässt sich ein Tensorprodukt endlich vieler Vektorräume V_1, \dots, V_k konstruieren, d.h. es existiert eine k -lineare Abbildung

$$V_1 \times \dots \times V_k \rightarrow V_1 \otimes \dots \otimes V_k, \quad (v_1, \dots, v_k) \mapsto v_1 \otimes \dots \otimes v_k,$$

mit folgender Eigenschaft: Zu jeder k -linearen Abbildung $\mu: V_1 \times \dots \times V_k \rightarrow W$ existiert eine eindeutige lineare Abbildung $\varphi: V_1 \otimes \dots \otimes V_k \rightarrow W$, sodass $\varphi(v_1 \otimes \dots \otimes v_k) = \mu(v_1, \dots, v_k)$, für alle $v_i \in V_i$. Durch diese universelle Eigenschaft ist das Tensorprodukt bis auf kanonischen Isomorphismus eindeutig bestimmt. Wir erhalten einen kanonischen linearen Isomorphismus

$$L(V_1, \dots, V_k; W) = L(V_1 \otimes \dots \otimes V_k, W).$$

Ist B_i eine Basis von V_i , dann bildet $\{b_1 \otimes \dots \otimes b_k : b_i \in B_i\}$ eine Basis von $V_1 \otimes \dots \otimes V_k$. Sind alle V_i endlich dimensional, dann ist auch ihr Tensorprodukt endlich dimensional, und es gilt

$$\dim(V_1 \otimes \dots \otimes V_k) = \dim(V_1) \cdots \dim(V_k).$$

Sind $\varphi_i: V_i \rightarrow W_i$ linear, dann existiert genau eine lineare Abbildung

$$V_1 \otimes \dots \otimes V_k \xrightarrow{\varphi_1 \otimes \dots \otimes \varphi_k} W_1 \otimes \dots \otimes W_k$$

sodass $(\varphi_1 \otimes \dots \otimes \varphi_k)(v_1, \dots, v_k) = \varphi_1(v_1) \cdots \varphi_k(v_k)$, für alle $v_i \in V_i$. Es gelten Rechenregeln analog zu denen in Proposition IX.2.7. Ist $\sigma \in \mathfrak{S}_k$ eine Permutation, so existiert eine eindeutige lineare Abbildung

$$\phi_\sigma: V_1 \otimes \dots \otimes V_k \rightarrow V_{\sigma(1)} \otimes \dots \otimes V_{\sigma(k)} \quad (\text{IX.8})$$

sodass $\phi_\sigma(v_1 \otimes \dots \otimes v_k) = v_{\sigma(1)} \otimes \dots \otimes v_{\sigma(k)}$. Aus der Eindeutigkeit erhalten wir

$$\phi_{\sigma_1 \sigma_2} = \phi_{\sigma_1} \circ \phi_{\sigma_2} \quad \text{und} \quad \phi_{\text{id}} = \text{id}_{V_1 \otimes \dots \otimes V_k}, \quad (\text{IX.9})$$

für beliebige Permutationen $\sigma_1, \sigma_2 \in \mathfrak{S}_k$. Daraus folgt auch, dass jedes ϕ_σ ein Isomorphismus mit Inverser $\phi_\sigma^{-1} = \phi_{\sigma^{-1}}$ ist. Insbesondere erhalten wir einen kanonischen linearen Isomorphismus

$$V \otimes W = W \otimes V, \quad v \otimes w \leftrightarrow w \otimes v.$$

Das Tensorprodukt ist in folgendem Sinn assoziativ:

IX.2.13. PROPOSITION. *Es existieren kanonische lineare Isomorphismen:*

$$V_1 \otimes (V_2 \otimes V_3) = V_1 \otimes V_2 \otimes V_3 = (V_1 \otimes V_2) \otimes V_3,$$

$$v_1 \otimes (v_2 \otimes v_3) \leftrightarrow v_1 \otimes v_2 \otimes v_3 \leftrightarrow (v_1 \otimes v_2) \otimes v_3.$$

BEWEIS. Aus (IX.2) und (IX.3) erhalten wir:

$$\begin{aligned} L_3(V_1, V_2, V_3; W) &= L(V_1, L_2(V_2, V_3; W)) \\ &= L(V_1, L(V_2 \otimes V_3, W)) = L(V_1 \otimes (V_2 \otimes V_3); W) \end{aligned}$$

Zu jeder trilinearen Abbildung $\mu: V_1 \times V_2 \times V_3 \rightarrow W$ existiert daher genau eine lineare Abbildung $\varphi: V_1 \otimes (V_2 \otimes V_3) \rightarrow W$, sodass $\varphi(v_1 \otimes (v_2 \otimes v_3)) = \mu(v_1, v_2, v_3)$. Somit hat die trilineare Abbildung

$$V_1 \times V_2 \times V_3 \rightarrow V_1 \otimes (V_2 \otimes V_3), \quad (v_1, v_2, v_3) \mapsto v_1 \otimes (v_2 \otimes v_3),$$

die universelle Eigenschaft des Tensorprodukts $V_1 \otimes V_2 \otimes V_3$, und wir erhalten den Isomorphismus $V_1 \otimes (V_2 \otimes V_3) = V_1 \otimes V_2 \otimes V_3$. Analog lässt sich der zweite Isomorphismus konstruieren. \square

IX.3. Erweiterung der Skalare. Jede lineare Abbildung zwischen reellen Vektorräumen kann in natürlicher Weise zu einer \mathbb{C} -linearen Abbildung zwischen komplexen Vektorräumen gemacht werden. Letztere sind oft leichter zu studieren, denn \mathbb{C} ist algebraisch abgeschlossen. Oft ist es möglich daraus Informationen über die ursprüngliche \mathbb{R} -lineare Abbildung zu erlangen. Bei der Behandlung der reellen Jordan'schen Normalform sind wir genau so vorgegangen. Im folgenden kurzen Abschnitt wollen wir diese Konstruktion mit Hilfe des Tensorprodukts systematischer untersuchen und verallgemeinern.

IX.3.1. DEFINITION (Teilkörper). Sei \mathbb{L} ein Körper. Unter einem *Teilkörper* auch *Unterkörper* von \mathbb{L} verstehen wir jede Teilmenge $\mathbb{K} \subseteq \mathbb{L}$, die bezüglich der Operationen von \mathbb{L} selbst einen Körper bildet. In diesem Fall kann \mathbb{L} als Vektorraum über \mathbb{K} aufgefasst werden. Die Dimension $\dim_{\mathbb{K}}(\mathbb{L})$ wird als Grad der Körpererweiterung $\mathbb{K} \subseteq \mathbb{L}$ bezeichnet.

IX.3.2. BEISPIEL. \mathbb{R} ist ein Teilkörper von \mathbb{C} . \mathbb{Q} ist ein Teilkörper von \mathbb{R} und auch Teilkörper von \mathbb{C} . Die Menge $\{a + b\sqrt{2} : a, b \in \mathbb{Q}\}$ bildet einen Teilkörper von \mathbb{R} . Die Menge $\{a + bi : a, b \in \mathbb{Q}\}$ bildet einen Teilkörper von \mathbb{C} .

Sei $\mathbb{K} \subseteq \mathbb{L}$ ein Teilkörper und V ein \mathbb{K} -Vektorraum. Es lässt sich dann $V \otimes_{\mathbb{K}} \mathbb{L}$ zu einem Vektorraum über \mathbb{L} machen, indem wir die Skalarmultiplikation mit $\lambda \in \mathbb{L}$ durch

$$V \otimes_{\mathbb{K}} \mathbb{L} \xrightarrow{\text{id}_V \otimes \lambda} V \otimes_{\mathbb{K}} \mathbb{L}, \quad \text{d.h.} \quad \lambda(v \otimes \mu) = v \otimes (\lambda\mu)$$

definieren, $v \in V$, $\lambda, \mu \in \mathbb{L}$. Dies ist wohldefiniert, denn die Skalarmultiplikation $\mathbb{L} \xrightarrow{\lambda} \mathbb{L}$ ist \mathbb{K} -linear. Es sind nun die Vektorraumaxiome zu überprüfen: Für $\lambda_1, \lambda_2, \mu \in \mathbb{L}$ und $v \in V$ gilt

$$\lambda_1(\lambda_2(v \otimes \mu)) = v \otimes (\lambda_1(\lambda_2\mu)) = v \otimes ((\lambda_1\lambda_2)\mu) = (\lambda_1\lambda_2)(v \otimes \mu)$$

und daher $\lambda_1(\lambda_2\xi) = (\lambda_1\lambda_2)\xi$, für alle $\xi \in V \otimes_{\mathbb{K}} \mathbb{L}$, wegen der Eindeutigkeitsaussage in der universellen Eigenschaft des Tensorprodukts.²⁰ Ebenso

$$\begin{aligned} (\lambda_1 + \lambda_2)(v \otimes \mu) &= v \otimes ((\lambda_1 + \lambda_2)\mu) = v \otimes (\lambda_1\mu + \lambda_2\mu) \\ &= v \otimes (\lambda_1\mu) + v \otimes (\lambda_2\mu) = \lambda_1(v \otimes \mu) + \lambda_2(v \otimes \mu), \end{aligned}$$

and daher $(\lambda_1 + \lambda_2)\xi = \lambda_1\xi + \lambda_2\xi$, für alle $\xi \in V \otimes_{\mathbb{K}} \mathbb{L}$, wieder aufgrund der Eindeutigkeitsaussage in der universellen Eigenschaft des Tensorprodukts. Alle anderen Vektorraumaxiome lassen sich noch einfacher verifizieren. Somit ist also $V \otimes_{\mathbb{K}} \mathbb{L}$ ein Vektorraum über \mathbb{L} .

Für eine lineare Abbildung $\varphi: V \rightarrow W$ zwischen \mathbb{K} -Vektorräumen, ist

$$V \otimes_{\mathbb{K}} \mathbb{L} \xrightarrow{\varphi \otimes \text{id}_{\mathbb{L}}} W \otimes_{\mathbb{K}} \mathbb{L}$$

eine \mathbb{L} -lineare Abbildung. Sind $v \in V$ und $\mu, \lambda \in \mathbb{L}$, dann gilt nämlich

$$\begin{aligned} (\varphi \otimes \text{id}_{\mathbb{L}})(\lambda(v \otimes \mu)) &= (\varphi \otimes \text{id}_{\mathbb{L}})(v \otimes (\lambda\mu)) = \varphi(v) \otimes (\lambda\mu) \\ &= \lambda(\varphi(v) \otimes \mu) = \lambda((\varphi \otimes \text{id}_{\mathbb{L}})(v \otimes \mu)), \end{aligned}$$

und daher $(\varphi \otimes \text{id}_{\mathbb{L}})(\lambda\xi) = \lambda((\varphi \otimes \text{id}_{\mathbb{L}})(\xi))$, für alle $\xi \in V \otimes_{\mathbb{K}} \mathbb{L}$. Beachte, dass diese Konstruktion mit der Komposition von Abbildungen verträglich ist, d.h.

$$(\psi \circ \varphi) \otimes \text{id}_{\mathbb{L}} = (\psi \otimes \text{id}_{\mathbb{L}}) \circ (\varphi \otimes \text{id}_{\mathbb{L}}) \quad \text{und} \quad \text{id}_V \otimes \text{id}_{\mathbb{L}} = \text{id}_{V \otimes_{\mathbb{K}} \mathbb{L}},$$

für jede weitere lineare Abbildung zwischen \mathbb{K} -Vektorräumen, $\psi: W \rightarrow U$.

Ist b_i , $i \in I$ eine Basis des \mathbb{K} -Vektorraums V , dann bildet $b_i \otimes 1$, $i \in I$ eine Basis des \mathbb{L} -Vektorraums $V \otimes_{\mathbb{K}} \mathbb{L}$. Nach Proposition IX.2.5 ist dieses System nämlich linear unabhängig, und jedes Element in $V \otimes_{\mathbb{K}} \mathbb{L}$ lässt sich in der Form $\sum_{i=1}^n b_i \otimes \lambda_i = \sum_{i=1}^n \lambda_i(b_i \otimes 1)$ schreiben, wobei $\lambda_i \in \mathbb{L}$. Insbesondere folgt

$$\dim_{\mathbb{L}}(V \otimes_{\mathbb{K}} \mathbb{L}) = \dim_{\mathbb{K}}(V).$$

Für jede lineare Abbildungen $\varphi: V \rightarrow W$ zwischen endlich dimensionalen \mathbb{K} -Vektorräumen gilt

$$\text{tr}_{\mathbb{L}}(\varphi \otimes \text{id}_{\mathbb{L}}) = \text{tr}_{\mathbb{K}}(\varphi) \quad \text{und} \quad \det_{\mathbb{L}}(\varphi \otimes \text{id}_{\mathbb{L}}) = \det_{\mathbb{K}}(\varphi)$$

siehe Aufgabe 59.

²⁰Beide Ausdrücke, $\lambda_1(\lambda_2\xi)$ und $(\lambda_1\lambda_2)\xi$ sind \mathbb{K} -linear in ξ .

Die \mathbb{K} -lineare und injektive Abbildung

$$\iota: V \rightarrow V \otimes_{\mathbb{K}} \mathbb{L}, \quad \iota(v) := v \otimes 1$$

hat folgender universelle Eigenschaft: *Ist W ein Vektorraum über \mathbb{L} , und ist $\varphi: V \rightarrow W$ eine \mathbb{K} -lineare Abbildung, dann existiert eine eindeutige \mathbb{L} -lineare Abbildung $\tilde{\varphi}: V \otimes_{\mathbb{K}} \mathbb{L} \rightarrow W$, sodass $\tilde{\varphi} \circ \iota = \varphi$:*

$$\begin{array}{ccc} V & \xrightarrow{\iota} & V \otimes_{\mathbb{K}} \mathbb{L} \\ & \searrow \varphi & \swarrow \exists! \tilde{\varphi} \\ & & W \end{array}$$

Aus der universellen Eigenschaft des Tensorprodukts $V \otimes_{\mathbb{K}} \mathbb{L}$ folgt nämlich, dass $\tilde{\varphi}(v \otimes \mu) := \mu\varphi(v)$ eine wohldefinierte \mathbb{K} -lineare Abbildung $V \otimes_{\mathbb{K}} \mathbb{L} \rightarrow W$ definiert, denn der Ausdruck $\mu\varphi(v)$ ist \mathbb{K} -linear in v und μ . Offensichtlich gilt $\tilde{\varphi} \circ \iota = \varphi$. Ist $\lambda \in \mathbb{L}$, dann $\tilde{\varphi}(\lambda(v \otimes \mu)) = \tilde{\varphi}(v \otimes (\lambda\mu)) = (\lambda\mu)\varphi(v) = \lambda(\mu\varphi(v)) = \lambda\tilde{\varphi}(v \otimes \mu)$, also $\tilde{\varphi}(\lambda\xi) = \lambda\tilde{\varphi}(\xi)$, für alle $\xi \in V \otimes_{\mathbb{K}} \mathbb{L}$, d.h. $\tilde{\varphi}$ ist auch \mathbb{L} -linear. Die universelle Eigenschaft liefert einen \mathbb{L} -linearen Isomorphismus

$$L_{\mathbb{L}}(V \otimes_{\mathbb{K}} \mathbb{L}, W) = L_{\mathbb{K}}(V, W_{\mathbb{K}}).$$

wobei $W_{\mathbb{K}}$ den W zugrundeliegenden \mathbb{K} -Vektorraum bezeichnet.

Ist V ein reeller Vektorraum, dann wird der komplexe Vektorraum $V \otimes_{\mathbb{R}} \mathbb{C}$ als *Komplexifizierung* von V bezeichnet. Bei der Behandlung der reellen Jordan'schen Normalform in Abschnitt VI.4 wurde diese erfolgreich verwendet, allerdings haben wir dort mit einer ad hoc Definition der Komplexifizierung gearbeitet.

IX.4. Tensoralgebra. Unter einer \mathbb{K} -Algebra verstehen wir einen \mathbb{K} -Vektorraum \mathcal{A} zusammen mit einer \mathbb{K} -bilinearen Abbildung, der *Multiplikation*,

$$\mathcal{A} \times \mathcal{A} \rightarrow \mathcal{A}, \quad (a, b) \mapsto ab.$$

Die Algebra wird *assoziativ* genannt, wenn $(ab)c = a(bc)$, für alle $a, b, c \in \mathcal{A}$. Ein Element $e \in \mathcal{A}$ heißt *Einselement*, falls $ea = a = ae$, für alle $a \in \mathcal{A}$. In diesem Fall ist das Einselement eindeutig bestimmt und wird meist mit $1 \in \mathcal{A}$ bezeichnet. Existieren Teilräume \mathcal{A}^q sodass $\mathcal{A} = \bigoplus_{q \in \mathbb{Z}} \mathcal{A}^q$, und gilt $ab \in \mathcal{A}^{p+q}$, für alle $a \in \mathcal{A}^p$ und $b \in \mathcal{A}^q$, dann sprechen wir von einer *graduerten Algebra*.

Seien \mathcal{A} und \mathcal{A}' zwei \mathbb{K} -Algebren. Unter einem *Algebrahomomorphismus* verstehen wir eine \mathbb{K} -lineare Abbildung $\varphi: \mathcal{A} \rightarrow \mathcal{A}'$, die auch mit der Multiplikation verträglich ist, d.h.

$$\varphi(ab) = \varphi(a)\varphi(b),$$

für alle $a, b \in \mathcal{A}$. Besitzen \mathcal{A} und \mathcal{A}' Einselemente, dann verlangen von einem Algebrahomomorphismus auch $\varphi(1) = 1$. Ein Algebrahomomorphismus $\varphi: \mathcal{A} \rightarrow \mathcal{A}'$ wird als *Algebraisomorphismus* bezeichnet, wenn ein Algebrahomomorphismus $\psi: \mathcal{A}' \rightarrow \mathcal{A}$ existiert, sodass $\psi \circ \varphi = \text{id}_{\mathcal{A}}$ und $\varphi \circ \psi = \text{id}_{\mathcal{A}'}$.

IX.4.1. BEISPIEL. Ist X eine Menge, dann bildet der Vektorraum aller Funktionen, $F(X, \mathbb{K})$, mit der üblichen (punktweisen) Multiplikation von Funktionen, eine \mathbb{K} -Algebra mit Eins. Ist X ein topologischer Raum, dann bildet die Menge aller stetigen Funktionen, $C(X; \mathbb{R})$, eine reelle Algebra. Analog ist die Menge der komplexwertigen stetigen Funktionen, $C(X; \mathbb{C})$, eine komplexe Algebra. Auch die Menge der stetigen Funktionen mit kompakten Träger, $C_c(X; \mathbb{R})$ bzw. $C_c(X; \mathbb{C})$, bilden Algebren über \mathbb{R} bzw. \mathbb{C} , besitzen jedoch für nicht kompaktes X kein Einselement. Alle diese Algebren sind kommutativ, d.h. es gilt $ab = ba$ für beliebige a und b .

IX.4.2. BEISPIEL. Ist V ein \mathbb{K} -Vektorraum, dann bildet $\text{end}(V)$ bezüglich der Komposition linearer Abbildungen eine, i.A. nicht kommutative Algebra. Ebenso ist $M_{n \times n}(\mathbb{K})$ bezüglich Matrizenmultiplikation eine nicht kommutative Algebra. Ist V ein n -dimensionaler Vektorraum, dann liefert jede Basis von V einen Algebraisomorphismus $M_{n \times n}(\mathbb{K}) \cong \text{end}(V)$.

Für jeden Vektorraum V und $q \in \mathbb{N}_0$ definieren wir

$$T^q V := \underbrace{V \otimes \cdots \otimes V}_{q \text{ Faktoren}}.$$

Insbesondere $T^0 V = \mathbb{K}$, $T^1 V = V$, $T^2 V = V \otimes V$ und $T^3 V = V \otimes V \otimes V$. Jede lineare Abbildung $\varphi: V \rightarrow W$ induziert eine lineare Abbildung

$$T^q \varphi: T^q V \rightarrow T^q W, \quad T^q \varphi := \varphi \otimes \cdots \otimes \varphi.$$

Offensichtlich gilt

$$T^q(\psi \circ \varphi) = (T^q \psi) \circ (T^q \varphi) \quad \text{und} \quad T^q \text{id}_V = \text{id}_{T^q V}, \quad (\text{IX.10})$$

für jede weitere lineare Abbildung $\psi: W \rightarrow U$. Darüber hinaus gilt, siehe (IX.8),

$$\phi_\sigma \circ T^q \varphi = (T^q \varphi) \circ \phi_\sigma, \quad (\text{IX.11})$$

für jede Permutation $\sigma \in \mathfrak{S}_q$, denn für $v_1, \dots, v_q \in V$ haben wir

$$\begin{aligned} \phi_\sigma((T^q \varphi)(v_1 \otimes \cdots \otimes v_q)) &= \phi_\sigma(\varphi(v_1) \otimes \cdots \otimes \varphi(v_q)) \\ &= \varphi(v_{\sigma(1)}) \otimes \cdots \otimes \varphi(v_{\sigma(q)}) \\ &= (T^q \varphi)(v_{\sigma(1)} \otimes \cdots \otimes v_{\sigma(q)}) \\ &= (T^q \varphi)(\phi_\sigma(v_1 \otimes \cdots \otimes v_q)), \end{aligned}$$

also auch $\phi_\sigma((T^q \varphi)\xi) = (T^q \varphi)(\phi_\sigma(\xi))$, für alle $\xi \in T^q V$.

Ist b_1, \dots, b_n eine Basis von V , dann bilden die Vektoren

$$b_{i_1} \otimes \cdots \otimes b_{i_q}, \quad 1 \leq i_1, \dots, i_q \leq n, \quad (\text{IX.12})$$

eine Basis von $T^q V$, siehe Proposition IX.2.5. Für jeden endlich dimensionalen Vektorraum V gilt daher

$$\dim(T^q V) = \dim(V)^q.$$

Wir definieren:

$$TV := \bigoplus_{q=0}^{\infty} T^q V = \mathbb{K} \oplus V \oplus (V \otimes V) \oplus (V \otimes V \otimes V) \oplus \dots$$

Beachte, dass TV unendlich dimensional ist, falls $V \neq 0$.

IX.4.3. SATZ (Tensoralgebra). Sei V ein Vektorraum über einem Körper \mathbb{K} . Die bilinearen Abbildungen $T^p V \times T^q V \xrightarrow{\otimes} T^{p+q} V$ induzieren eine bilineare Abbildung

$$TV \times TV \xrightarrow{\otimes} TV.$$

Dadurch wird TV zu einer assoziativen graduierten Algebra mit Einselement $1 \in \mathbb{K} = T^0 V \subseteq TV$. Jede lineare Abbildung $\varphi: V \rightarrow W$ induziert einen Algebromorphismus $T\varphi = \bigoplus_{q=0}^{\infty} T^q \varphi: TV \rightarrow TW$, und es gilt

$$T(\psi \circ \varphi) = (T\psi) \circ (T\varphi) \quad \text{sowie} \quad T(\text{id}_V) = \text{id}_{TV},$$

für jede weitere lineare Abbildung $\psi: W \rightarrow U$. Die kanonische lineare Inklusion $\iota: V \rightarrow TV$, die wir aus der Identifikation $V = T^1 V \subseteq TV$ erhalten, hat folgende universelle Eigenschaft: Ist \mathcal{A} eine assoziative \mathbb{K} -Algebra mit Eins und $\varphi: V \rightarrow \mathcal{A}$ linear, dann existiert genau ein Algebromorphismus $\tilde{\varphi}: TV \rightarrow \mathcal{A}$, sodass $\tilde{\varphi} \circ \iota = \varphi$, d.h. $\tilde{\varphi}|_V = \varphi$:

$$\begin{array}{ccc} V & \xrightarrow{\iota} & TV \\ & \searrow \varphi & \swarrow \tilde{\varphi} \\ & & \mathcal{A} \end{array}$$

Durch diese Eigenschaft ist TV zusammen mit der Abbildung $\iota: V \rightarrow TV$ bis auf kanonischen Algebriomorphismus eindeutig bestimmt, d.h. ist \tilde{T} eine weitere assoziative \mathbb{K} -Algebra mit Eins und $\tilde{\iota}: V \rightarrow \tilde{T}$ eine lineare Abbildung, die die selben Eigenschaft wie $\iota: V \rightarrow TV$ hat, dann existiert ein eindeutiger Algebriomorphismus $\phi: TV \xrightarrow{\cong} \tilde{T}$, sodass $\phi \circ \iota = \tilde{\iota}$.

BEWEIS. Aus den vorangehenden Betrachtungen ist klar, dass TV eine graduierte assoziative Algebra mit Eins bildet. Wir werden daher nur die universelle Eigenschaft der Abbildung $\iota: V \rightarrow TV$ verifizieren. Sei also $\varphi: V \rightarrow \mathcal{A}$ linear. Dann existiert eine eindeutige lineare Abbildung $\tilde{\varphi}^p: T^p V \rightarrow \mathcal{A}$, sodass

$$\tilde{\varphi}^p(v_1 \otimes \dots \otimes v_p) = \varphi(v_1) \cdots \varphi(v_p), \quad (\text{IX.13})$$

für alle $v_i \in V$, denn die rechte Seite ist linear in jedem v_i . Diese induzieren eine lineare Abbildung $\tilde{\varphi}: TV \rightarrow \mathcal{A}$, sodass $\tilde{\varphi}|_{T^p V} = \tilde{\varphi}^p$. Offensichtlich ist $\tilde{\varphi}$ ein Algebromorphismus und es gilt $\tilde{\varphi} \circ \iota = \varphi$. Auch die Eindeutigkeit von $\tilde{\varphi}$ ist offensichtlich, denn für einen Algebromorphismus muss (IX.13) gelten.

Aus der universellen Eigenschaft von TV erhalten wir einen eindeutigen Algebromorphismus $\phi: TV \rightarrow \tilde{T}$, sodass $\phi \circ \iota = \tilde{\iota}$. Aus dem gleichen Grund existiert ein Algebromorphismus $\psi: \tilde{T} \rightarrow TV$, sodass $\psi \circ \tilde{\iota} = \iota$. Wir erhalten

$\psi \circ \phi \circ \iota = \iota = \text{id}_{TV} \circ \iota$, also $\psi \circ \phi = \text{id}_{TV}$ wegen der Eindeutigkeitsaussage in der universellen Eigenschaft von TV . Analog lässt sich $\phi \circ \psi = \text{id}_{\bar{T}}$ zeigen. Folglich sind ϕ und ψ zueinander inverse Algebrasomorphismen. \square

IX.5. Grassmann Algebra. Wir beschränken uns in diesem Abschnitt auf Vektorräume über einem Körper \mathbb{K} mit Charakteristik 0. Sei V ein Vektorraum über \mathbb{K} . Wir erinnern uns, dass jede Permutation $\sigma \in \mathfrak{S}_q$ einen linearen Isomorphismus

$$\phi_\sigma: T^q V \rightarrow T^q V, \quad \phi_\sigma(v_1 \otimes \cdots \otimes v_q) = v_{\sigma(1)} \otimes \cdots \otimes v_{\sigma(q)},$$

induziert, sodass $\phi_{\sigma\tau} = \phi_\sigma \circ \phi_\tau$ und $\phi_{\text{id}} = \text{id}_{T^q V}$, für aller $\sigma, \tau \in \mathfrak{S}_q$. Es bezeichne

$$\Lambda^q V := \{ \xi \in T^q V \mid \forall \sigma \in \mathfrak{S}_q : \phi_\sigma(\xi) = \text{sign}(\sigma)\xi \},$$

den Teilraum der *alternierenden* Tensoren. Beachte, $\Lambda^0 V = \mathbb{K}$ und $\Lambda^1 V = V$. Ist $\varphi: V \rightarrow W$ eine lineare Abbildung, dann folgt aus (IX.11), dass die lineare Abbildung $T^q \varphi: T^q V \rightarrow T^q W$ den Teilraum $\Lambda^q V$ in den Teilraum $\Lambda^q W$ abbildet. Wir bezeichnen diese Einschränkung mit

$$\Lambda^q \varphi: \Lambda^q V \rightarrow \Lambda^q W.$$

Aus (IX.10) erhalten wir sofort

$$\Lambda^q(\psi \circ \varphi) = (\Lambda^q \psi) \circ (\Lambda^q \varphi) \quad \text{und} \quad \Lambda^q(\text{id}_V) = \text{id}_{\Lambda^q V},$$

für jede weitere lineare Abbildung $\psi: W \rightarrow U$.

IX.5.1. LEMMA. *Die lineare Abbildung*

$$A: T^q V \rightarrow T^q V, \quad A := \frac{1}{q!} \sum_{\sigma \in \mathfrak{S}_q} \text{sign}(\sigma) \phi_\sigma,$$

ist ein Projektor, $A^2 = A$, mit $\text{img}(A) = \Lambda^q V$, und es gilt

$$A \circ \phi_\sigma = \text{sign}(\sigma)A = \phi_\sigma \circ A, \tag{IX.14}$$

für jede Permutation $\sigma \in \mathfrak{S}_q$.

BEWEIS. Wir beginnen mit (IX.14): Da \mathfrak{S}_q aus $q!$ Permutationen besteht:

$$\begin{aligned} A \circ \phi_\tau &= \frac{1}{q!} \sum_{\sigma \in \mathfrak{S}_q} \text{sign}(\sigma) \phi_\sigma \circ \phi_\tau = \frac{1}{q!} \sum_{\sigma \in \mathfrak{S}_q} \text{sign}(\sigma) \phi_{\sigma\tau} \\ &= \frac{1}{q!} \sum_{\sigma \in \mathfrak{S}_q} \text{sign}(\sigma\tau^{-1}) \phi_\sigma = \text{sign}(\tau) \frac{1}{q!} \sum_{\sigma \in \mathfrak{S}_q} \text{sign}(\sigma) \phi_\sigma = \text{sign}(\tau)A. \end{aligned}$$

Analog lässt sich die zweite Gleichheit in (IX.14) zeigen. Daraus folgt nun:

$$A \circ A = \frac{1}{q!} \sum_{\sigma \in \mathfrak{S}_q} \text{sign}(\sigma) \phi_\sigma \circ A = \frac{1}{q!} \sum_{\sigma \in \mathfrak{S}_q} A = A.$$

Somit ist A ein Projektor, aus (IX.14) folgt sofort $\text{img}(A) = \Lambda^q V$. \square

Für zwei Vektorräume V und W bezeichne $L_q^{\text{alt}}(V; W) \subseteq L_q(V, \dots, V; W)$ den Vektorraum der q -linearen alternierenden Abbildungen $\mu: V \times \dots \times V \rightarrow W$. Diese können durch jede der folgenden äquivalenten Eigenschaften charakterisiert werden:

- (a) $\mu(v_{\sigma(1)}, \dots, v_{\sigma(q)}) = \text{sign}(\sigma)\mu(v_1, \dots, v_q)$, für alle $v_1, \dots, v_q \in V$ und $\sigma \in \mathfrak{S}_q$.
- (b) $\mu(v_1, \dots, v_q)$ wechselt das Vorzeichen wenn zwei Eintragungen v_i und v_j vertauscht werden, $i \neq j$.
- (c) $\mu(v_1, \dots, v_q) = 0$, falls v_1, \dots, v_q linear abhängig sind.
- (d) $\mu(v_1, \dots, v_q) = 0$, falls $i \neq j$ existiert, sodass $v_i = v_j$.
- (e) $\mu(v_1, \dots, v_q) = 0$, falls i existiert, sodass $v_i = v_{i+1}$.

IX.5.2. PROPOSITION. *Die q -lineare alternierende Abbildung*

$$V \times \dots \times V \rightarrow \Lambda^q V, \quad (v_1, \dots, v_q) \mapsto v_1 \wedge \dots \wedge v_q := A(v_1 \otimes \dots \otimes v_q),$$

besitzt folgende universelle Eigenschaft: Ist $\mu: V \times \dots \times V \rightarrow U$ eine q -lineare alternierende Abbildung, dann existiert genau eine lineare Abbildung $\varphi: \Lambda^q V \rightarrow U$, sodass $\varphi(v_1 \wedge \dots \wedge v_q) = \mu(v_1, \dots, v_q)$, für alle $v_1, \dots, v_q \in V$:

$$\begin{array}{ccc} V \times \dots \times V & \xrightarrow{\quad} & \Lambda^q V \\ & \searrow \mu & \swarrow \exists! \varphi \\ & & U \end{array}$$

Dies liefert einen kanonischen linearen Isomorphismus

$$L_q^{\text{alt}}(V; U) = L(\Lambda^q V, U).$$

Durch obige universelle Eigenschaft ist $\Lambda^q V$ zusammen mit der q -linearen alternierenden Abbildung $V \times \dots \times V \rightarrow \Lambda^q V$ bis auf kanonischen Isomorphismus eindeutig bestimmt.

BEWEIS. Aus (IX.14) folgt sofort, dass $A(v_1 \otimes \dots \otimes v_q)$ alternierend ist. Da $\{v_1 \otimes \dots \otimes v_q : v_1, \dots, v_q \in V\}$ ein Erzeugendensystem von $T^q V$ bildet, ist $\{A(v_1 \otimes \dots \otimes v_q) : v_1, \dots, v_q \in V\}$ ein Erzeugendensystem von $\Lambda^q V = \text{img}(A)$. Daraus folgt sofort, dass die lineare Abbildung $\varphi: \Lambda^q V \rightarrow U$ durch die Bedingung $\varphi(v_1 \wedge \dots \wedge v_q) = \mu(v_1, \dots, v_q)$ eindeutig bestimmt ist, falls sie existiert. Um die Existenz zu zeigen, bezeichne $\tilde{\varphi}: T^q V \rightarrow U$ die eindeutige lineare Abbildung, sodass

$$\tilde{\varphi}(v_1 \otimes \dots \otimes v_q) = \mu(v_1, \dots, v_q),$$

für alle $v_1, \dots, v_q \in V$. Da μ alternierend ist, gilt

$$\begin{aligned} (\tilde{\varphi} \circ \phi_\sigma)(v_1 \otimes \dots \otimes v_q) &= \tilde{\varphi}(v_{\sigma(1)} \otimes \dots \otimes v_{\sigma(q)}) = \mu(v_{\sigma(1)}, \dots, v_{\sigma(q)}) \\ &= \text{sign}(\sigma)\mu(v_1, \dots, v_q) = \text{sign}(\sigma)\tilde{\varphi}(v_1 \otimes \dots \otimes v_q), \end{aligned}$$

für jede Permutation $\sigma \in \mathfrak{S}_q$, also

$$\tilde{\varphi} \circ \phi_\sigma = \text{sign}(\sigma)\tilde{\varphi}.$$

Daraus erhalten wir $\tilde{\varphi} \circ A = \tilde{\varphi}$, denn

$$\tilde{\varphi} \circ A = \frac{1}{q!} \sum_{\sigma \in \mathfrak{S}_q} \text{sign}(\sigma) \tilde{\varphi} \circ \phi_\sigma = \frac{1}{q!} \sum_{\sigma \in \mathfrak{S}_q} \tilde{\varphi} = \tilde{\varphi}.$$

Bezeichnet nun $\varphi := \tilde{\varphi}|_{\Lambda^q V} : \Lambda^q V \rightarrow U$ die Einschränkung, so folgt

$$\varphi(v_1 \wedge \cdots \wedge v_q) = \tilde{\varphi}(A(v_1 \otimes \cdots \otimes v_q)) = \tilde{\varphi}(v_1 \otimes \cdots \otimes v_q) = \mu(v_1, \dots, v_q).$$

Die verbleibenden Behauptungen sind nun offensichtlich. \square

IX.5.3. BEISPIEL. Die Determinant kann als

$$\det \in L_n^{\text{alt}}(\mathbb{K}^n; \mathbb{K}) = L(\Lambda^n \mathbb{K}^n, \mathbb{K})$$

aufgefasst werden.

Wir definieren:

$$\Lambda V := \bigoplus_{q=0}^{\infty} \Lambda^q V = \underbrace{\Lambda^0 V}_{=\mathbb{K}} \oplus \underbrace{\Lambda^1 V}_{=V} \oplus \Lambda^2 V \oplus \Lambda^3 V \oplus \cdots$$

IX.5.4. SATZ (Grassmann Algebra). Die bilineare Abbildung

$$\Lambda^p V \times \Lambda^q V \xrightarrow{\wedge} \Lambda^{p+q} V, \quad (\alpha, \beta) \mapsto \alpha \wedge \beta := A(\alpha \otimes \beta) \quad (\text{IX.15})$$

hat folgende Eigenschaften:

- (a) $\alpha \wedge (\beta \wedge \gamma) = (\alpha \wedge \beta) \wedge \gamma$, für alle $\alpha \in \Lambda^p V$, $\beta \in \Lambda^q V$ und $\gamma \in \Lambda^r V$.
- (b) $\alpha \wedge \beta = (-1)^{pq} \beta \wedge \alpha$, für alle $\alpha \in \Lambda^p V$ und $\beta \in \Lambda^q V$.
- (c) $(\Lambda^{p+q} \varphi)(\alpha \wedge \beta) = (\Lambda^p \varphi)(\alpha) \wedge (\Lambda^q \varphi)(\beta)$ für jede lineare Abbildung $\varphi : V \rightarrow W$.

D.h. die von (IX.15) induzierte bilineare Abbildung $\Lambda V \times \Lambda V \xrightarrow{\wedge} \Lambda V$ macht ΛV zu einer graduiert kommutativen und assoziativen Algebra mit Einselement $1 \in \mathbb{K} = \Lambda^0 V \subseteq \Lambda V$. Diese Algebra wird als äußere Algebra oder Grassmann Algebra bezeichnet. Jede lineare Abbildung $\varphi : V \rightarrow W$ induziert einen Algebrahomomorphismus $\Lambda \varphi : \Lambda V \rightarrow \Lambda W$, $\Lambda \varphi = \bigoplus_{q=0}^{\infty} \Lambda^q \varphi$, und es gilt

$$\Lambda(\psi \circ \varphi) = (\Lambda \psi) \circ (\Lambda \varphi) \quad \text{sowie} \quad (\Lambda \text{id}_V) = \text{id}_{\Lambda V},$$

für jede weitere lineare Abbildung $\psi : W \rightarrow U$. Die lineare und injektive Abbildung $\iota : V \rightarrow \Lambda V$, die wir aus der Identifikation $V = \Lambda^1 V \subseteq \Lambda V$ erhalten besitzt folgende universelle Eigenschaft: Ist \mathcal{A} eine assoziative \mathbb{K} -Algebra mit Eins, und $\varphi : V \rightarrow \mathcal{A}$ eine lineare Abbildung, sodass $\varphi(v)\varphi(v) = 0$, für alle $v \in V$, dann existiert ein eindeutig bestimmter Algebrahomomorphismus $\tilde{\varphi} : \Lambda V \rightarrow \mathcal{A}$, sodass $\tilde{\varphi} \circ \iota = \varphi$, d.h. $\tilde{\varphi}|_V = \varphi$:

$$\begin{array}{ccc} V & \xrightarrow{\iota} & \Lambda V \\ & \searrow \varphi & \swarrow \exists! \tilde{\varphi} \\ & & \mathcal{A} \end{array}$$

Durch diese universelle Eigenschaft ist ΛV zusammen mit der Abbildung ι bis auf kanonischen Isomorphismus eindeutig bestimmt.

BEWEIS. Wir beginnen mit (b). Es bezeichne $\tau \in \mathfrak{S}_{p+q}$ folgende Permutation:

$$\tau(i) = \begin{cases} i+q & \text{falls } 1 \leq i \leq p \\ i-p & \text{falls } p+1 \leq i \leq p+q. \end{cases}$$

Für beliebige $v_i \in V$ gilt daher:

$$\phi_\tau(v_1 \otimes \cdots \otimes v_p \otimes v_{p+1} \otimes \cdots \otimes v_{p+q}) = v_{p+1} \otimes \cdots \otimes v_{p+q} \otimes v_1 \otimes \cdots \otimes v_p$$

also auch $\phi_\tau(\xi \otimes \eta) = \eta \otimes \xi$, für alle $\xi \in T^pV$ und $\eta \in T^qV$. Insbesondere

$$\beta \otimes \alpha = \phi_\tau(\alpha \otimes \beta).$$

Da $\text{sign}(\tau) = (-1)^{pq}$ folgt mit (IX.14)

$$\beta \wedge \alpha = A(\beta \otimes \alpha) = A(\phi_\tau(\alpha \otimes \beta)) = \text{sign}(\tau)A(\alpha \otimes \beta) = (-1)^{pq}\alpha \wedge \beta.$$

Um auch (a) zu zeigen, werden wir unten

$$A((A\xi) \otimes \eta) = A(\xi \otimes \eta) = A(\xi \otimes A\eta) \quad (\text{IX.16})$$

für alle $\xi \in T^pV$ und $\eta \in T^qV$ herleiten. Ist dies gelungen, dann folgt

$$\begin{aligned} \alpha \wedge (\beta \wedge \gamma) &= A(\alpha \otimes A(\beta \otimes \gamma)) \\ &= A(\alpha \otimes \beta \otimes \gamma) = A(A(\alpha \otimes \beta) \otimes \gamma) = (\alpha \wedge \beta) \wedge \gamma. \end{aligned}$$

Um (IX.16) zu zeigen, fassen wir Permutationen $\sigma \in \mathfrak{S}_p$ auch als Permutationen $\tilde{\sigma} \in \mathfrak{S}_{p+q}$ auf:

$$\tilde{\sigma}(i) := \begin{cases} \sigma(i) & \text{falls } 1 \leq i \leq p \\ i & \text{falls } p+1 \leq i \leq p+q. \end{cases}$$

Es gilt daher $(\phi_\sigma\xi) \otimes \eta = \phi_{\tilde{\sigma}}(\xi \otimes \eta)$, für alle $\xi \in T^pV$ und $\eta \in T^qV$. Da offensichtlich $\text{sign}(\tilde{\sigma}) = \text{sign}(\sigma)$ erhalten wir aus (IX.14)

$$A((\phi_\sigma\xi) \otimes \eta) = A(\phi_{\tilde{\sigma}}(\xi \otimes \eta)) = \text{sign}(\tilde{\sigma})A(\xi \otimes \eta) = \text{sign}(\sigma)A(\xi \otimes \eta)$$

und somit

$$A((A\xi) \otimes \eta) = \frac{1}{p!} \sum_{\sigma \in \mathfrak{S}_p} \text{sign}(\sigma)A((\phi_\sigma\xi) \otimes \eta) = \frac{1}{p!} \sum_{\sigma \in \mathfrak{S}_p} A(\xi \otimes \eta) = A(\xi \otimes \eta).$$

Die zweite Gleichheit in (IX.16) lässt sich völlig analog behandeln. Um (c) zu zeigen, verwenden wir (IX.11) um

$$(\Lambda^q\varphi) \circ A = A \circ (T^q\varphi)$$

herzuleiten. Es folgt dann:

$$\begin{aligned} (\Lambda^{p+q}\varphi)(\alpha \wedge \beta) &= (\Lambda^{p+q}\varphi)(A(\alpha \otimes \beta)) \\ &= A((T^{p+q}\varphi)(\alpha \otimes \beta)) = A((T^p\varphi)(\alpha) \otimes (T^q\varphi)(\beta)) \\ &= A((\Lambda^p\varphi)(\alpha) \otimes (\Lambda^q\varphi)(\beta)) = (\Lambda^p\varphi)(\alpha) \wedge (\Lambda^q\varphi)(\beta). \end{aligned}$$

Nun zur universellen Eigenschaft der Abbildung $\iota: V \rightarrow \Lambda V$. Nach Proposition IX.5.2 existiert eine eindeutige lineare Abbildung $\tilde{\varphi}^q: \Lambda^q V \rightarrow \mathcal{A}$, sodass

$$\tilde{\varphi}^q(v_1 \wedge \cdots \wedge v_q) = \varphi(v_1) \cdots \varphi(v_q),$$

denn die rechte Seite ist nach Voraussetzung alternierend. Diese Abbildungen induzieren eine lineare Abbildung $\tilde{\varphi}: \Lambda V = \bigoplus_q \Lambda^q V \rightarrow \mathcal{A}$, für die offensichtlich $\tilde{\varphi} \circ \iota = \varphi$ gilt. Es lässt sich leicht zeigen, dass dies ein Algebromorphismus ist, der durch $\tilde{\varphi} \circ \iota = \varphi$ völlig festgelegt ist. Die verbleibenden Behauptungen sind nun offensichtlich. \square

IX.5.5. PROPOSITION. *Sei V ein n -dimensionaler Vektorraum und b_1, \dots, b_n eine Basis von V . Dann bilden die Vektoren*

$$b_{i_1} \wedge \cdots \wedge b_{i_q}, \quad 1 \leq i_1 < i_2 < \cdots < i_q \leq n, \quad (\text{IX.17})$$

eine Basis von $\Lambda^q V$. Insbesondere gilt $\Lambda^q V = 0$, für alle $q > n$ und daher

$$\Lambda V = \bigoplus_{q=0}^n \Lambda^q V = \Lambda^0 V \oplus \Lambda^1 V \oplus \Lambda^2 V \oplus \cdots \oplus \Lambda^n V.$$

Weiters gilt

$$\dim(\Lambda^q V) = \binom{n}{q} \quad \text{und} \quad \dim(\Lambda V) = 2^n.$$

BEWEIS. Da die Vektoren $b_{i_1} \otimes \cdots \otimes b_{i_q}$, $1 \leq i_1, \dots, i_q \leq n$, eine Basis von $T^q V$ bilden, ist

$$b_{i_1} \wedge \cdots \wedge b_{i_q} = A(b_{i_1} \otimes \cdots \otimes b_{i_q}), \quad 1 \leq i_1, \dots, i_q \leq n,$$

ein Erzeugendensystem von $\Lambda^q V = \text{img}(A)$. Nach Satz IX.5.4(b) haben wir weiters $b_{\sigma(i_1)} \wedge \cdots \wedge b_{\sigma(i_q)} = \pm b_{i_1} \wedge \cdots \wedge b_{i_q}$, für jede Permutation $\sigma \in \mathfrak{S}_q$, und es gilt $b_{i_1} \wedge \cdots \wedge b_{i_q} = 0$, falls wenigstens zwei der Indizes i_1, \dots, i_q übereinstimmen. Somit bilden auch die Vektoren

$$b_{i_1} \wedge \cdots \wedge b_{i_q}, \quad 1 \leq i_1 < \cdots < i_q \leq n,$$

ein Erzeugendensystem von $\Lambda^q V$.

Um die lineare Unabhängigkeit einzusehen, seien $\lambda_{i_1 \dots i_q} \in \mathbb{K}$, sodass

$$\sum_{1 \leq i_1 < \cdots < i_q \leq n} \lambda_{i_1 \dots i_q} b_{i_1} \wedge \cdots \wedge b_{i_q} = 0.$$

Da $b_{i_1} \wedge \cdots \wedge b_{i_q} = A(b_{i_1} \otimes \cdots \otimes b_{i_q})$, folgt

$$\begin{aligned} 0 &= A\left(\sum_{1 \leq i_1 < \cdots < i_q \leq n} \lambda_{i_1 \dots i_q} b_{i_1} \otimes \cdots \otimes b_{i_q}\right) \\ &= \sum_{\sigma \in \mathfrak{S}_q} \sum_{1 \leq i_1 < \cdots < i_q \leq n} \frac{\text{sign}(\sigma) \lambda_{i_1 \dots i_q}}{q!} b_{\sigma(i_1)} \otimes \cdots \otimes b_{\sigma(i_q)}. \end{aligned}$$

Beachte, dass die Vektoren

$$b_{\sigma(i_1)} \otimes \cdots \otimes b_{\sigma(i_q)}, \quad \sigma \in \mathfrak{S}_q, \quad 1 \leq i_1 < \cdots < i_q \leq n$$

linear unabhängig in $T^q V$ sind, siehe (IX.12). Folglich $\frac{\text{sign}(\sigma)\lambda_{i_1 \dots i_q}}{q!} = 0$, also $\lambda_{i_1 \dots i_q} = 0$, für alle $1 \leq i_1 < \cdots < i_q \leq n$. Dies zeigt, dass die Vektoren (IX.17) auch linear unabhängig, also eine Basis von $\Lambda^q V$ sind. Da diese Basis aus $\binom{n}{q}$ Elementen besteht, folgt $\dim(\Lambda^q V) = \binom{n}{q}$ und dann $\dim(\Lambda V) = \sum_{q=0}^n \dim(\Lambda^q V) = \sum_{q=0}^n \binom{n}{q} = (1+1)^n = 2^n$ mit Hilfe des binomischen Lehrsatzes. \square

IX.5.6. BEISPIEL. Sei V ein n -dimensionaler Vektorraum und $\varphi: V \rightarrow V$ linear. Nach Proposition IX.5.5 ist $\Lambda^n V$ ein 1-dimensionaler Vektorraum, die lineare Abbildung $\Lambda^n \varphi: \Lambda^n V \rightarrow \Lambda^n V$ ist daher durch Multiplikation mit einem Skalar gegeben. Tatsächlich gilt

$$\Lambda^n \varphi = \det(\varphi) \text{id}_{\Lambda^n V}.$$

Um dies einzusehen, nehmen wir zunächst an, dass φ diagonalisierbar ist. Es existiert daher eine Basis b_1, \dots, b_n von V und Skalare $\lambda_i \in \mathbb{K}$ mit $\varphi(b_i) = \lambda_i b_i$. Nach (IX.17) bildet der Vektor $b_1 \wedge \cdots \wedge b_n$ eine Basis von $\Lambda^n V$ und

$$\begin{aligned} (\Lambda^n \varphi)(b_1 \wedge \cdots \wedge b_n) &= \varphi(b_1) \wedge \cdots \wedge \varphi(b_n) \\ &= \lambda_1 b_1 \wedge \cdots \wedge \lambda_n b_n = \det(\varphi) b_1 \wedge \cdots \wedge b_n. \end{aligned}$$

Der allgemein Fall ist in Aufgabe 64 zu verifizieren.

IX.5.7. BEISPIEL. Sei V ein n -dimensionaler Vektorraum und $\varphi: V \rightarrow V$ linear. Dann gilt für die Koeffizienten des charakteristischen Polynoms:

$$\det(\varphi + z \text{id}_V) = \sum_{q=0}^n \text{tr}(\Lambda^q \varphi) z^{n-q}.$$

Wir nehmen wieder an, dass φ diagonalisierbar ist. Es existiert daher eine Basis b_1, \dots, b_n von V und Skalare $\lambda_i \in \mathbb{K}$ mit $\varphi(b_i) = \lambda_i b_i$. Somit

$$(\Lambda^q \varphi)(b_{i_1} \wedge \cdots \wedge b_{i_q}) = \varphi(b_{i_1}) \wedge \cdots \wedge \varphi(b_{i_q}) = \lambda_{i_1} \cdots \lambda_{i_q} b_{i_1} \wedge \cdots \wedge b_{i_q}.$$

Aus (IX.17) folgt daher

$$\text{tr}(\Lambda^q \varphi) = \sum_{1 \leq i_1 < \cdots < i_q \leq n} \lambda_{i_1} \cdots \lambda_{i_q},$$

also

$$\begin{aligned} \det(\varphi + z \text{id}_V) &= \prod_{j=1}^n (\lambda_j + z) \\ &= \sum_{q=0}^n \sum_{1 \leq i_1 < \cdots < i_q \leq n} \lambda_{i_1} \cdots \lambda_{i_q} z^{n-q} = \sum_{q=0}^n \text{tr}(\Lambda^q \varphi) z^{n-q}. \end{aligned}$$

Der allgemeine Fall ist in Aufgabe 65 zu behandeln.

IX.5.8. BEISPIEL. Seien V und W zwei endlich dimensionale Vektorräume. Weiters bezeichnen $\iota_1: V \rightarrow V \oplus W$ und $\iota_2: W \rightarrow V \oplus W$ die beiden kanonischen Inklusionen. Dann induzieren die linearen Abbildungen

$$\Lambda^p V \otimes \Lambda^q W \rightarrow \Lambda^{p+q}(V \oplus W), \quad \alpha \otimes \beta \mapsto (\Lambda^p \iota_1)(\alpha) \wedge (\Lambda^q \iota_2)(\beta),$$

eine lineare Abbildung $\bigoplus_{p+q=k} \Lambda^p V \otimes \Lambda^q W \rightarrow \Lambda^k(V \oplus W)$. Aus Proposition IX.5.5 folgt, dass diese Basen auf Basen abbildet. Wir erhalten daher einen kanonischen linearen Isomorphismus

$$\Lambda^k(V \oplus W) = \bigoplus_{p+q=k} \Lambda^p V \otimes \Lambda^q W$$

und auch $\Lambda(V \oplus W) = (\Lambda V) \otimes (\Lambda W)$.

IX.5.9. BEISPIEL. Die Komposition $\Lambda^q V^* \subseteq T^q V^* = (T^q V)^* \rightarrow (\Lambda^q V)^*$, siehe (IX.6), liefert einen kanonischen linearen Isomorphismus

$$\Lambda^q V^* = (\Lambda^q V)^* = L_q^{\text{alt}}(V; \mathbb{K}).$$

Für $\alpha \in \Lambda^p V^* = L_p^{\text{alt}}(V; \mathbb{K})$ und $\beta \in \Lambda^q V^* = L_q^{\text{alt}}(V; \mathbb{K})$ ist $\alpha \wedge \beta \in \Lambda^{p+q} V^* = L_{p+q}^{\text{alt}}(V; \mathbb{K})$ durch

$$\begin{aligned} & (\alpha \wedge \beta)(v_1, \dots, v_{p+q}) \\ &= \frac{1}{(p+q)!} \sum_{\sigma \in \mathfrak{S}_{p+q}} \text{sign}(\sigma) \alpha(v_{\sigma(1)}, \dots, v_{\sigma(p)}) \beta(v_{\sigma(p+1)}, \dots, v_{\sigma(p+q)}) \end{aligned}$$

gegeben, wobei $v_1, \dots, v_{p+q} \in V$.

Die äußere Algebra ΛV kann auch als Quotient der Tensoralgebra TV definiert werden, und diese Konstruktion ist für Körper beliebiger Charakteristik sinnvoll. Details dazu finden sich in [11, Chapter 4.6].

IX.6. Symmetrische Algebra. Wir beschränken uns auch in diesem Abschnitt auf Vektorräume über einem Körper \mathbb{K} der Charakteristik 0. Sei V ein Vektorraum über \mathbb{K} . Es bezeichne

$$S^q V := \{ \xi \in T^q V \mid \forall \sigma \in \mathfrak{S}_q : \phi_\sigma(\xi) = \xi \},$$

den Teilraum der *symmetrischen* Tensoren. Beachte, $S^0 V = \mathbb{K}$ und $S^1 V = V$. Ist $\varphi: V \rightarrow W$ eine lineare Abbildung, dann folgt aus (IX.11), dass die lineare Abbildung $T^q \varphi: T^q V \rightarrow T^q W$ den Teilraum $S^q V$ in den Teilraum $S^q W$ abbildet. Wir bezeichnen diese Einschränkung mit

$$S^q \varphi: S^q V \rightarrow S^q W.$$

Aus (IX.10) erhalten wir sofort

$$S^q(\psi \circ \varphi) = (S^q \psi) \circ (S^q \varphi) \quad \text{und} \quad S^q(\text{id}_V) = \text{id}_{S^q V},$$

für jede weitere lineare Abbildung $\psi: W \rightarrow U$.

IX.6.1. LEMMA. Die lineare Abbildung

$$S: T^q V \rightarrow T^q V, \quad S(\xi) := \frac{1}{q!} \sum_{\sigma \in \mathfrak{S}_q} \phi_\sigma(\xi),$$

ist ein Projektor, $S^2 = S$, mit $\text{img}(S) = S^q V$, und es gilt

$$S \circ \phi_\sigma = S = \phi_\sigma \circ S, \quad (\text{IX.18})$$

für jede Permutation $\sigma \in \mathfrak{S}_q$.

BEWEIS. Wir beginnen mit (IX.18): Da \mathfrak{S}_q aus $q!$ Permutationen besteht:

$$S \circ \phi_\tau = \frac{1}{q!} \sum_{\sigma \in \mathfrak{S}_q} \phi_\sigma \circ \phi_\tau = \frac{1}{q!} \sum_{\sigma \in \mathfrak{S}_q} \phi_{\sigma\tau} = \frac{1}{q!} \sum_{\sigma \in \mathfrak{S}_q} \phi_\sigma = S.$$

Analog lässt sich die zweite Gleichheit in (IX.18) zeigen. Daraus folgt nun:

$$S \circ S = \frac{1}{q!} \sum_{\sigma \in \mathfrak{S}_q} \phi_\sigma \circ S = \frac{1}{q!} \sum_{\sigma \in \mathfrak{S}_q} S = S.$$

Somit ist S ein Projektor, aus (IX.18) folgt sofort $\text{img}(S) = S^q V$. \square

Für zwei Vektorräume V und W bezeichne $L_q^{\text{sym}}(V; W) \subseteq L_q(V, \dots, V; W)$ den Vektorraum der q -linearen symmetrischen Abbildungen $\mu: V \times \dots \times V \rightarrow W$. Diese können durch jede der folgenden äquivalenten Eigenschaften charakterisiert werden:

- (a) $\mu(v_{\sigma(1)}, \dots, v_{\sigma(q)}) = \mu(v_1, \dots, v_q)$, für alle $v_1, \dots, v_q \in V$ und $\sigma \in \mathfrak{S}_q$.
- (b) $\mu(v_1, \dots, v_q)$ bleibt unverändert wenn zwei Eintragungen vertauscht werden.
- (c) $\mu(v_1, \dots, v_q)$, bleibt unverändert wenn zwei benachbarte Eintragungen vertauscht werden.

IX.6.2. PROPOSITION. Die q -lineare symmetrische Abbildung

$$V \times \dots \times V \rightarrow S^q V, \quad (v_1, \dots, v_q) \mapsto v_1 \cdots v_q := S(v_1 \otimes \cdots \otimes v_q),$$

besitzt folgende universelle Eigenschaft: Ist $\mu: V \times \dots \times V \rightarrow U$ eine q -lineare symmetrische Abbildung, dann existiert genau eine lineare Abbildung $\varphi: S^q V \rightarrow U$, sodass $\varphi(v_1 \cdots v_q) = \mu(v_1, \dots, v_q)$, für alle $v_1, \dots, v_q \in V$:

$$\begin{array}{ccc} V \times \dots \times V & \xrightarrow{\quad} & S^q V \\ & \searrow \mu & \swarrow \exists! \varphi \\ & & U \end{array}$$

Dies liefert einen kanonischen linearen Isomorphismus

$$L_q^{\text{sym}}(V; U) = L(S^q V, U).$$

Durch obige universelle Eigenschaft ist $S^q V$ zusammen mit der q -linearen symmetrischen Abbildung $V \times \dots \times V \rightarrow S^q V$ bis auf kanonischen Isomorphismus eindeutig bestimmt.

BEWEIS. Aus (IX.18) folgt sofort, dass $S(v_1 \otimes \cdots \otimes v_q)$ symmetrisch ist. Da $\{v_1 \otimes \cdots \otimes v_q : v_1, \dots, v_q \in V\}$ ein Erzeugendensystem von $T^q V$ bildet, ist $\{S(v_1 \otimes \cdots \otimes v_q) : v_1, \dots, v_q \in V\}$ ein Erzeugendensystem von $S^q V = \text{img}(S)$. Daraus folgt sofort, dass die lineare Abbildung $\varphi: S^q V \rightarrow U$ durch die Bedingung $\varphi(v_1 \cdots v_q) = \mu(v_1, \dots, v_q)$ eindeutig bestimmt ist, falls sie existiert. Um die Existenz zu zeigen, bezeichne $\tilde{\varphi}: T^q V \rightarrow U$ die eindeutige lineare Abbildung, sodass

$$\tilde{\varphi}(v_1 \otimes \cdots \otimes v_q) = \mu(v_1, \dots, v_q),$$

für alle $v_1, \dots, v_q \in V$. Da μ symmetrisch ist, gilt

$$\begin{aligned} (\tilde{\varphi} \circ \phi_\sigma)(v_1 \otimes \cdots \otimes v_q) &= \tilde{\varphi}(v_{\sigma(1)} \otimes \cdots \otimes v_{\sigma(q)}) = \mu(v_{\sigma(1)}, \dots, v_{\sigma(q)}) \\ &= \mu(v_1, \dots, v_q) = \tilde{\varphi}(v_1 \otimes \cdots \otimes v_q), \end{aligned}$$

für jede Permutation $\sigma \in \mathfrak{S}_q$, also

$$\tilde{\varphi} \circ \phi_\sigma = \tilde{\varphi}.$$

Daraus erhalten wir $\tilde{\varphi} \circ S = \tilde{\varphi}$, denn

$$\tilde{\varphi} \circ S = \frac{1}{q!} \sum_{\sigma \in \mathfrak{S}_q} \tilde{\varphi} \circ \phi_\sigma = \frac{1}{q!} \sum_{\sigma \in \mathfrak{S}_q} \tilde{\varphi} = \tilde{\varphi}.$$

Bezeichnet nun $\varphi := \tilde{\varphi}|_{S^q V}: S^q V \rightarrow U$ die Einschränkung, so folgt

$$\varphi(v_1 \cdots v_q) = \tilde{\varphi}(S(v_1 \otimes \cdots \otimes v_q)) = \tilde{\varphi}(v_1 \otimes \cdots \otimes v_q) = \mu(v_1, \dots, v_q).$$

Die verbleibenden Behauptungen sind nun offensichtlich. \square

Wir definieren:

$$SV := \bigoplus_{q=0}^{\infty} S^q V = \underbrace{S^0 V}_{=\mathbb{K}} \oplus \underbrace{S^1 V}_{=V} \oplus S^2 V \oplus S^3 V \oplus \cdots$$

Genau wie in Satz IX.5.4 lässt sich zeigen:

IX.6.3. SATZ (Symmetrische Algebra). *Die bilineare Abbildung*

$$S^p V \times S^q V \rightarrow S^{p+q} V, \quad (\alpha, \beta) \mapsto \alpha\beta := S(\alpha \otimes \beta) \quad (\text{IX.19})$$

hat folgende Eigenschaften:

(a) $\alpha(\beta\gamma) = (\alpha\beta)\gamma$, für alle $\alpha \in S^p V$, $\beta \in S^q V$ und $\gamma \in S^r V$.

(b) $\alpha\beta = \beta\alpha$, für alle $\alpha \in S^p V$ und $\beta \in S^q V$.

(c) $(S^{p+q}\varphi)(\alpha\beta) = (S^p\varphi)(\alpha)(S^q\varphi)(\beta)$ für jede lineare Abbildung $\varphi: V \rightarrow W$.

D.h. die von (IX.19) induzierte bilineare Abbildung $SV \times SV \rightarrow SV$ macht SV zu einer kommutativen und assoziativen graduierten Algebra mit Einselement $1 \in \mathbb{K} = S^0 V \subseteq SV$. Diese Algebra wird als symmetrische Algebra bezeichnet. Jede lineare Abbildung $\varphi: V \rightarrow W$ induziert einen Algebromorphismus $S\varphi: SV \rightarrow SW$, $S\varphi = \bigoplus_{q=0}^{\infty} S^q \varphi$, und es gilt

$$S(\psi \circ \varphi) = (S\psi) \circ (S\varphi) \quad \text{sowie} \quad S(\text{id}_V) = \text{id}_{SV},$$

für jede weitere lineare Abbildung $\psi: W \rightarrow U$. Die lineare und injektive Abbildung $\iota: V \rightarrow SV$, die wir aus der Identifikation $V = S^1V \subseteq SV$ erhalten besitzt folgende universelle Eigenschaft: Ist \mathcal{A} eine kommutative und assoziative \mathbb{K} -Algebra mit Eins, und $\varphi: V \rightarrow \mathcal{A}$ eine lineare Abbildung, dann existiert ein eindeutig bestimmter Algebromorphismus $\tilde{\varphi}: SV \rightarrow \mathcal{A}$, sodass $\tilde{\varphi} \circ \iota = \varphi$, d.h. $\tilde{\varphi}|_V = \varphi$:

$$\begin{array}{ccc} V & \xrightarrow{\iota} & SV \\ & \searrow \varphi & \swarrow \exists! \tilde{\varphi} \\ & & \mathcal{A} \end{array}$$

Durch diese universelle Eigenschaft ist SV zusammen mit der Abbildung ι bis auf kanonischen Isomorphismus eindeutig bestimmt.

IX.6.4. PROPOSITION. Sei V ein n -dimensionaler Vektorraum und b_1, \dots, b_n eine Basis von V . Dann bilden die Vektoren²¹

$$b_1^{a_1} \cdots b_n^{a_n}, \quad a_1, \dots, a_n \in \mathbb{N}_0, \quad a_1 + \cdots + a_n = q, \quad (\text{IX.20})$$

eine Basis von S^qV . Insbesondere gilt

$$\dim(S^qV) = \binom{n+q-1}{q}.$$

BEWEIS. Wie zuvor, lässt sich leicht zeigen, dass die Vektoren (IX.20) ein Erzeugendensystem von S^qV bilden. Um auch die lineare Unabhängigkeit einzusehen, seien $\lambda_{a_1 \dots a_n} \in \mathbb{K}$, sodass

$$\sum_{a_1 + \cdots + a_n = q} \lambda_{a_1 \dots a_n} b_1^{a_1} \cdots b_n^{a_n} = 0.$$

Dann gilt also

$$S \left(\sum_{a_1 + \cdots + a_n = q} \lambda_{a_1 \dots a_n} \underbrace{b_1 \otimes \cdots \otimes b_1}_{a_1 \text{ Faktoren}} \otimes \cdots \otimes \underbrace{b_n \otimes \cdots \otimes b_n}_{a_n \text{ Faktoren}} \right) = 0.$$

Daraus lässt sich $\lambda_{a_1 \dots a_n} = 0$ herleiten. Dies zeigt, dass die Vektoren (IX.20) linear unabhängig in S^qV sind. Da die Basis (IX.20) aus $\binom{n+q-1}{q}$ Elementen besteht, folgt $\dim(S^qV) = \binom{n+q-1}{q}$. \square

IX.6.5. BEISPIEL. Seien V und W zwei endlich dimensionale Vektorräume. Weiters bezeichnen $\iota_1: V \rightarrow V \oplus W$ und $\iota_2: W \rightarrow V \oplus W$ die beiden kanonischen Inklusionen. Dann induzieren die linearen Abbildungen

$$S^pV \otimes S^qW \rightarrow S^{p+q}(V \oplus W), \quad \alpha \otimes \beta \mapsto (S^p\iota_1)(\alpha)(S^q\iota_2)(\beta),$$

²¹Da es nicht auf die Reihenfolge ankommt schreiben wir hier

$$b_1^{a_1} \cdots b_n^{a_n} := \underbrace{b_1 \cdots b_1}_{a_1 \text{ Faktoren}} \cdots \underbrace{b_n \cdots b_n}_{a_n \text{ Faktoren}},$$

wobei $b_i^0 := 1$.

eine lineare Abbildung $\bigoplus_{p+q=k} S^p V \otimes S^q W \rightarrow S^k(V \oplus W)$. Aus Proposition IX.6.4 folgt, dass diese Basen auf Basen abbildet. Wir erhalten daher einen kanonischen linearen Isomorphismus

$$S^k(V \oplus W) = \bigoplus_{p+q=k} S^p V \otimes S^q W$$

und auch $S(V \oplus W) = (SV) \otimes (SW)$.

IX.6.6. BEISPIEL (Polynomialalgebra). Die Komposition

$$S^q V^* \subseteq T^q V^* = (T^q V)^* \rightarrow (S^q V)^*,$$

siehe (IX.6), liefert einen kanonischen linearen Isomorphismus

$$S^q V^* = (S^q V)^* = L_q^{\text{sym}}(V; \mathbb{K}).$$

Mittels Polarisieren, lässt sich zeigen, dass Elemente $\mu \in L_q^{\text{sym}}(V; \mathbb{K})$ durch die Werte $\mu(v) := \mu(v, \dots, v)$ eindeutig bestimmt sind. Für $\alpha \in S^p V^* = L_p^{\text{sym}}(V; \mathbb{K})$ und $\beta \in S^q V^* = L_q^{\text{sym}}(V; \mathbb{K})$ ist $\alpha\beta \in S^{p+q} V^* = L_{p+q}^{\text{sym}}(V; \mathbb{K})$ durch

$$(\alpha\beta)(v) = \alpha(v)\beta(v),$$

d.h. durch punktweise Multiplikation gegeben. Somit kann SV^* mit der Algebra der Polynome $V \rightarrow \mathbb{K}$ identifiziert werden.

Die symmetrische Algebra SV kann auch als Quotient der Tensoralgebra TV definiert werden, und diese Konstruktion ist für Körper beliebiger Charakteristik sinnvoll. Details dazu finden sich in [11, Chapter 4.5].

Literatur

- [1] A. Čap, Einführung in die Lineare Algebra und Geometrie. Skriptum zur gleichnamigen Vorlesung aus dem Sommersemester 2010. Erhältlich unter <http://www.mat.univie.ac.at/~cap/files/Linalg.pdf>
- [2] W. H. Greub, *Linear algebra*. Third edition. Die Grundlehren der Mathematischen Wissenschaften **97**. Springer-Verlag, New York, 1967.
- [3] K. Jänich, *Lineare Algebra*. Springer Verlag, 2008.
- [4] K. Jänich, *Topologie*. Springer Verlag, Berlin–Heidelberg, 1990.
- [5] S. Lang, *Linear algebra*. Undergraduate Texts in Mathematics, Springer-Verlag, New York, 1987.
- [6] S. Roman, *Advanced linear algebra*. Graduate Texts in Mathematics **135**, Springer, New York, 2008.
- [7] H. Schichl und R. Steinbauer, *Einführung in das mathematische Arbeiten*. Springer Verlag, 2009.
- [8] G. Strang, *Linear algebra and its applications*. Academic Press, New York–London, 1980.
- [9] A. Čap, *Lineare Algebra und Geometrie für LAK*. Skriptum zur gleichnamigen Vorlesung aus dem Wintersemester 2010/11. Erhältlich unter <http://www.mat.univie.ac.at/~cap/files/LinalgLAK.pdf>
- [10] K. Hoffman and R. Kunze, *Linear algebra*. Second edition. Prentice-Hall, 1971.
- [11] A.I. Kostrikin and Y.I. Manin, *Linear Algebra and Geometry*. Gordon and Breach Science Publishers, 1997.
- [12] W. Klingenberg, *Lineare Algebra und Geometrie*. Springer-Verlag, 1984.

Übungen zu “Einführung in die lineare Algebra und Geometrie”

Wintersemester 2011/12

Stefan Haller

1. Seien $n, m \in \mathbb{N}$ und $a_{ij} \in \mathbb{R}$. Zeige, dass die Abbildung $\psi: \mathbb{R}^n \rightarrow \mathbb{R}^m$,

$$\psi \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} := \begin{pmatrix} a_{11}x_1 + \cdots + a_{1n}x_n \\ \vdots \\ a_{m1}x_1 + \cdots + a_{mn}x_n \end{pmatrix} = x_1 \begin{pmatrix} a_{11} \\ \vdots \\ a_{m1} \end{pmatrix} + \cdots + x_n \begin{pmatrix} a_{1n} \\ \vdots \\ a_{mn} \end{pmatrix},$$

linear ist, d.h. für alle $x, \tilde{x} \in \mathbb{R}^n$ und alle $\lambda \in \mathbb{R}$ gilt

$$\psi(x + \tilde{x}) = \psi(x) + \psi(\tilde{x}) \quad \text{und} \quad \psi(\lambda x) = \lambda \psi(x).$$

2. Sei $\psi: \mathbb{R}^n \rightarrow \mathbb{R}^m$ eine lineare Abbildung, d.h. es gelte $\psi(x + \tilde{x}) = \psi(x) + \psi(\tilde{x})$ und $\psi(\lambda x) = \lambda \psi(x)$, für alle $x, \tilde{x} \in \mathbb{R}^n$ und alle $\lambda \in \mathbb{R}$. Zeige, dass die Menge

$$L := \{x \in \mathbb{R}^n \mid \psi(x) = 0\}$$

einen Teilraum bildet, d.h. für alle $x, \tilde{x} \in L$ und $\lambda \in \mathbb{R}$ gilt $x + \tilde{x} \in L$ und $\lambda x \in L$. Zeige weiters, dass auch

$$W := \text{img}(\psi) = \{y \in \mathbb{R}^m \mid \exists x \in \mathbb{R}^n : \psi(x) = y\}$$

ein Teilraum ist, d.h. für alle $y, \tilde{y} \in W$ und alle $\lambda \in \mathbb{R}$ gilt $y + \tilde{y} \in W$ und $\lambda y \in W$.

3. Zeige, dass das Gleichungssystem

$$\begin{array}{cccccc} x_1 & +2x_2 & +3x_3 & & = & y_1 \\ 2x_1 & +5x_2 & +8x_3 & +3x_4 & = & y_2 \\ 3x_1 & +8x_2 & +14x_3 & +8x_4 & = & y_3 \\ & +3x_2 & +8x_3 & +14x_4 & = & y_4 \end{array}$$

für jedes $y \in \mathbb{R}^4$ genau eine Lösung $x \in \mathbb{R}^4$ besitzt und bestimme diese Lösung. Was bedeutet dies für die Abbildung

$$\psi: \mathbb{R}^4 \rightarrow \mathbb{R}^4, \quad \psi \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} := \begin{pmatrix} x_1 + 2x_2 + 3x_3 \\ 2x_1 + 5x_2 + 8x_3 + 3x_4 \\ 3x_1 + 8x_2 + 14x_3 + 8x_4 \\ +3x_2 + 8x_3 + 14x_4 \end{pmatrix}?$$

4. Für welche $y \in \mathbb{R}^4$ besitzt das Gleichungssystem

$$\begin{array}{cccccc} 2x_1 & -3x_2 & 2x_3 & = & y_1 \\ 4x_1 & -6x_2 & -4x_3 & = & y_2 \\ 6x_1 & -9x_2 & -6x_3 & = & y_3 \\ -2x_1 & +3x_2 & +2x_3 & = & y_4 \end{array}$$

wenigstens eine Lösung $x \in \mathbb{R}^3$. Gib ein Gleichungssystem mit möglichst wenigen Gleichungen für die Menge dieser y an.

5. Betrachte das homogene Gleichungssystem

$$\begin{array}{cccccc} x_1 & & & +4x_4 & +5x_5 & = & 0 \\ x_1 & +2x_2 & & +10x_4 & +13x_5 & = & 0 \\ -x_1 & +x_2 & +3x_3 & +5x_4 & +8x_5 & = & 0 \\ 2x_1 & +2x_2 & +x_3 & +16x_4 & +21x_5 & = & 0 \end{array}$$

Bestimme eine Basis des Lösungsraums

$$L := \{x \in \mathbb{R}^5 \mid x \text{ genügen allen vier Gleichungen oben}\},$$

d.h. bestimme Vektoren $b_1, \dots, b_l \in L$, sodass sich jedes $x \in L$ in der Form

$$x = s_1 b_1 + \dots + s_l b_l$$

schreiben lässt, für eindeutig bestimmte Skalare $s_1, \dots, s_l \in \mathbb{R}$. *Hinweis:* $l = 2$. Gib auch ein Gleichungssystem für L an, das nur aus drei Gleichungen besteht.

6. Sei $\psi: \mathbb{R}^n \rightarrow \mathbb{R}^m$ linear, d.h. es gelte $\psi(x + \tilde{x}) = \psi(x) + \psi(\tilde{x})$ und $\psi(\lambda x) = \lambda\psi(x)$, für alle $x, \tilde{x} \in \mathbb{R}^n$ und alle $\lambda \in \mathbb{R}$. Für $y \in \mathbb{R}^m$ betrachte

$$L_y := \{x \in \mathbb{R}^n \mid \psi(x) = y\}$$

und setze $L := L_0 = \{x \in \mathbb{R}^n \mid \psi(x) = 0\}$. Weiters sei $\xi_y \in L_y$. Zeige, dass die Abbildung

$$\phi_y: L \xrightarrow{\cong} L_y, \quad \phi_y(x) := \xi_y + x,$$

eine Bijektion ist und schließe daraus

$$L_y = \{\xi_y + x \mid \psi(x) = 0\} = \xi_y + L.$$

Was bedeutet dies für ein lineares Gleichungssystem mit m Gleichungen in n Variablen?

7. Sei $\psi: \mathbb{R}^n \rightarrow \mathbb{R}^m$ linear und $\xi \in \mathbb{R}^m$. Zeige, dass die Abbildung

$$\phi: \mathbb{R}^n \rightarrow \mathbb{R}^m, \quad \phi(x) := \xi + \psi(x),$$

affin ist, d.h. es gilt

$$\phi(\lambda x + (1 - \lambda)\tilde{x}) = \lambda\phi(x) + (1 - \lambda)\phi(\tilde{x}),$$

für alle $x, \tilde{x} \in \mathbb{R}^n$ und alle $\lambda \in \mathbb{R}$. Wann ist ϕ linear?

8. Bestimme alle komplexen Lösungen des Gleichungssystems

$$\begin{array}{cccccc} z_1 & & +\mathbf{i}z_2 & +(1 + \mathbf{i})z_3 & = & \mathbf{i} - 1 \\ (2 + \mathbf{i})z_1 & +(-3 + \mathbf{i})z_2 & & +2\mathbf{i}z_3 & = & -5 \end{array}$$

d.h. bestimme alle $z = (z_1, z_2, z_3) \in \mathbb{C}^3$ die den beiden Gleichungen genügen.

9. Zeige, dass das Gleichungssystem

$$\begin{aligned} \mathbf{i}z_1 &+ 2\mathbf{i}z_2 &+ (1 + \mathbf{i})z_3 &= w_1 \\ (1 + \mathbf{i})z_1 &+ (2 + \mathbf{i})z_2 &+ (3 + 2\mathbf{i})z_3 &= w_2 \\ (2 + \mathbf{i})z_1 &+ (4 + 2\mathbf{i})z_2 &+ (1 + 2\mathbf{i})z_3 &= w_3 \end{aligned}$$

für jedes $w \in \mathbb{C}^3$ eine eindeutige Lösung $z \in \mathbb{C}^3$ besitzt und bestimme diese Lösung. Was bedeutet dies für die Abbildung

$$\psi: \mathbb{C}^3 \rightarrow \mathbb{C}^3, \quad \psi \begin{pmatrix} z_1 \\ z_2 \\ z_3 \end{pmatrix} := \begin{pmatrix} \mathbf{i}z_1 + 2\mathbf{i}z_2 + (1 + \mathbf{i})z_3 \\ (1 + \mathbf{i})z_1 + (2 + \mathbf{i})z_2 + (3 + 2\mathbf{i})z_3 \\ (2 + \mathbf{i})z_1 + (4 + 2\mathbf{i})z_2 + (1 + 2\mathbf{i})z_3 \end{pmatrix}?$$

10. Bestimme alle rationalen Lösungen des Gleichungssystems

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

d.h. bestimme alle $x = (x_1, x_2, x_3) \in \mathbb{Q}^3$, die dieses System lösen.

11. Sei \mathbb{K} ein Körper und $n \in \mathbb{N}$. Zeige, dass \mathbb{K}^n bezüglich komponentenweiser Addition und Skalarmultiplikation einen \mathbb{K} -Vektorraum bildet, vgl. Beispiel II.1.3 aus der Vorlesung.

12. Sei $\mathbb{K} = \mathbb{Z}_p$, wobei p eine Primzahl bezeichnet. Aus wievielen Elementen besteht der Vektorraum \mathbb{K}^n ?

13. Sei X eine Menge und V ein Vektorraum über \mathbb{K} . Zeige, dass die Menge aller Abbildungen $X \rightarrow V$ bezüglich punktweiser Addition und Skalarmultiplikation einen \mathbb{K} -Vektorraum bildet. *Hinweis: In der Vorlesung haben wir bereits die Vektorraumaxiome (V1) – (V3) und (V6) überprüft, siehe Beispiel II.1.6. Zeige nun analog, dass auch die restlichen Axiome gelten.*

14. Sei X eine Menge und \mathbb{K} ein Körper. Zeige, dass die punktweise Multiplikation \mathbb{K} -wertiger Funktionen $X \rightarrow \mathbb{K}$ folgende Eigenschaften besitzt:

- (a) $f(gh) = (fg)h$
- (b) $fg = gf$
- (c) $1f = f = f1$, wobei $1: X \rightarrow \mathbb{K}$ die konstante Einsfunktion bezeichnet.
- (d) $f(g_1 + g_2) = fg_1 + fg_2$ und $f(\lambda g) = \lambda(fg)$
- (e) $(f_1 + f_2)g = f_1g + f_2g$ und $(\lambda f)g = \lambda(fg)$

für beliebige Funktionen $f, f_1, f_2, g, g_1, g_2, h: X \rightarrow \mathbb{K}$ und $\lambda \in \mathbb{K}$. Die Menge aller Funktion $X \rightarrow \mathbb{K}$ bildet daher eine kommutative \mathbb{K} -Algebra mit Eins, vgl. Bemerkung II.1.7 aus der Vorlesung.

15. Zeige, dass $\mathbb{K}[z]$, d.h. die Menge der Polynome mit Koeffizienten in einem Körper \mathbb{K} , tatsächlich einen \mathbb{K} -Vektorraum bildet. Zeige weiters, dass die Multiplikation von Polynomen folgende Eigenschaften besitzt:

- (a) $p(qr) = (pq)r$

(b) $pq = qp$

(c) $1p = p = p1$, wobei $1 = 1 + 0z + 0z^2 + \dots$ das Einspolynom bezeichnet.

(d) $r(p + q) = rp + rq$ und $p(\lambda q) = \lambda(pq)$

(e) $(p + q)r = pr + qr$ und $(\lambda p)q = \lambda(pq)$

für beliebige Polynome $p, q, r \in \mathbb{K}[z]$ und $\lambda \in \mathbb{K}$. Somit ist $\mathbb{K}[z]$ eine kommutative \mathbb{K} -Algebra mit Eins. *Hinweis: In der Vorlesung haben wir bereits die Assoziativität bewiesen, vgl. Bemerkung II.1.9.*

16. Welche der folgenden Teilmengen sind Teilräume von \mathbb{R}^2 ?

$$\begin{array}{ll} \left\{ \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2 : 2x + 3y = 7 \right\} & \left\{ \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2 : x = 0 \right\} \\ \left\{ \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2 : y = x^2 \right\} & \left\{ \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2 : x^2 = 0 \right\} \\ \left\{ \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2 : x \leq y \right\} & \left\{ \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2 : x + y = 0 \text{ und } x - y = 0 \right\} \end{array}$$

17. Wieviele Elemente hat der Teilraum

$$\left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{K}^3 \mid x + y + z = 0 \right\},$$

über dem Körper $\mathbb{K} = \mathbb{Z}_2$ und wieviele im Fall $\mathbb{K} = \mathbb{Z}_3$?

18. Zeige, dass die Menge der 1-periodischen Funktionen,

$$\{f: \mathbb{R} \rightarrow \mathbb{R} \mid \forall x \in \mathbb{R} : f(x+1) = f(x)\},$$

einen Teilraum von $F(\mathbb{R}, \mathbb{R})$ bildet.

19. Seien W_1 und W_2 zwei Teilräume eines Vektorraums V . Zeige, dass die Vereinigung $W_1 \cup W_2$ genau dann einen Teilraum von V bildet, wenn $W_1 \subseteq W_2$ oder $W_2 \subseteq W_1$ gilt.

20. Sei $W_i, i \in I$, eine Familie von Teilräumen eines Vektorraums V , sodass für je zwei $i, j \in I$ stets $W_i \subseteq W_j$ oder $W_j \subseteq W_i$ gilt. Zeige, dass in dieser Situation $\bigcup_{i \in I} W_i$ einen Teilraum von V bildet.

21. Welche der folgenden Abbildungen sind linear?

$$\varphi: \mathbb{R}^2 \rightarrow \mathbb{R}, \quad \varphi \begin{pmatrix} x \\ y \end{pmatrix} = 0 \qquad \psi: \mathbb{R}^3 \rightarrow \mathbb{R}^3, \quad \psi(v) = -v$$

$$\rho: \mathbb{R}^2 \rightarrow \mathbb{R}^2, \quad \rho \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} xy \\ 2x \end{pmatrix} \qquad \kappa: \mathbb{R}^2 \rightarrow \mathbb{R}^2, \quad \kappa \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x^3 \\ y^3 \end{pmatrix}$$

$$\chi: \mathbb{R}^3 \rightarrow \mathbb{R}, \quad \chi \begin{pmatrix} x \\ y \\ z \end{pmatrix} = x - y + z \qquad f: \mathbb{R} \rightarrow \mathbb{R}, \quad f(x) = 3x + 7$$

22. Sei $\varphi: V \rightarrow W$ eine Abbildung zwischen \mathbb{K} -Vektorräumen. Zeige, dass φ genau dann linear ist, wenn es folgender Bedingung genügt:

$$\forall \lambda \in \mathbb{K} \forall v_1, v_2 \in V : \varphi(v_1 + \lambda v_2) = \varphi(v_1) + \lambda \varphi(v_2)$$

23. Sei $\varphi: V \rightarrow W$ eine Abbildung zwischen \mathbb{Q} -Vektorräumen. Zeige, dass φ genau dann linear ist, wenn es folgende Eigenschaft besitzt:

$$\forall v, w \in V : \varphi(v + w) = \varphi(v) + \varphi(w)$$

Hinweis: Zeige zunächst $\varphi(nv) = n\varphi(v)$, für alle $n \in \mathbb{N}$ und $v \in V$.

24. Sei V ein Vektorraum und $g: Y \rightarrow X$ eine Abbildung. Zeige, dass die Zuordnung $F(X, V) \rightarrow F(Y, V)$, $f \mapsto f \circ g$, eine lineare Abbildung ist. Schließe daraus, dass für jede Teilmenge $A \subseteq X$, die Abbildung $F(X, V) \rightarrow F(A, V)$, $f \mapsto f|_A$, linear ist. Folgere daraus auch, dass für jedes $x \in X$, die sogenannte Evaluationsabbildung, $\text{ev}_x: F(X, V) \rightarrow V$, $\text{ev}_x(f) := f(x)$, linear ist.

25. Sei \mathbb{K} ein Körper. Betrachte die beiden Abbildungen

$$\varphi: \mathbb{K}^2 \rightarrow \mathbb{K}^2, \varphi \begin{pmatrix} x \\ y \end{pmatrix} := \begin{pmatrix} x + y \\ y \end{pmatrix}, \quad \text{und} \quad \psi: \mathbb{K}^2 \rightarrow \mathbb{K}^2, \psi \begin{pmatrix} x \\ y \end{pmatrix} := \begin{pmatrix} x \\ x + y \end{pmatrix}.$$

Zeige, dass φ und ψ beide lineare Isomorphismen sind, und bestimme ihre Umkehrabbildungen. Zeige auch $\varphi \circ \psi \neq \psi \circ \varphi$, und schließe daraus, dass die Gruppe $\text{GL}(\mathbb{K}^2)$ nicht abelsch ist.

26. Sei \mathbb{K} ein Körper und betrachte den Teilraum

$$W := \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{K}^3 \mid 7x - y + 8z = 0 \right\}$$

von \mathbb{K}^3 . Konstruiere einen linearen Isomorphismus $\mathbb{K}^2 \cong W$.

27. Sei \mathbb{K} ein Körper und betrachte den Teilraum

$$W := \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{K}^3 \mid \begin{array}{l} x + 2y + 3z = 0 \\ 2x + 5y + z = 0 \end{array} \right\}$$

von \mathbb{K}^3 . Konstruiere einen linearen Isomorphismus $\mathbb{K}^2 \cong W$.

28. Sei W ein Teilraum eines Vektorraums V . Zeige, dass

$$\{\varphi \in \text{end}(V) \mid \varphi(W) \subseteq W\}$$

einen Teilraum von $\text{end}(V)$ bildet, der auch abgeschlossen unter der Komposition linearer Abbildungen ist. Zeige auch, dass

$$\{\varphi \in \text{GL}(V) \mid \varphi(W) = W\}$$

eine Untergruppe von $\text{GL}(V)$ bildet.

29 (Fünferlemma). Freiwilliges Beispiel! Sei \mathbb{K} ein Körper und

$$\begin{array}{ccccccccc}
 V_1 & \xrightarrow{\varphi_1} & V_2 & \xrightarrow{\varphi_2} & V_3 & \xrightarrow{\varphi_3} & V_4 & \xrightarrow{\varphi_4} & V_5 \\
 \rho_1 \downarrow \cong & & \rho_2 \downarrow \cong & & \downarrow \rho_3 & \cong & \downarrow \rho_4 & \cong & \downarrow \rho_5 \\
 W_1 & \xrightarrow{\psi_1} & W_2 & \xrightarrow{\psi_2} & W_3 & \xrightarrow{\psi_3} & W_4 & \xrightarrow{\psi_4} & W_5
 \end{array}$$

ein kommutatives Diagramm linearer Abbildungen zwischen \mathbb{K} -Vektorräumen mit exakten Zeilen. Genauer, sollen die folgenden Voraussetzungen erfüllt sein:

- (a) V_1, \dots, V_5 und W_1, \dots, W_5 sind \mathbb{K} -Vektorräume.
- (b) $\varphi_1, \dots, \varphi_4, \psi_1, \dots, \psi_4$ und ρ_1, \dots, ρ_5 sind lineare Abbildungen.
- (c) Das Diagramm kommutiert, d.h. für $i = 1, 2, 3, 4$ gilt

$$\rho_{i+1} \circ \varphi_i = \psi_i \circ \rho_i.$$

- (d) Die Zeilen sind exakt, d.h. für $i = 1, 2, 3$ gilt

$$\text{img}(\varphi_i) = \ker(\varphi_{i+1}) \quad \text{und} \quad \text{img}(\psi_i) = \ker(\psi_{i+1}).$$

Zeige nun: Sind ρ_1, ρ_2, ρ_4 und ρ_5 Isomorphismen, dann muss auch ρ_3 ein Isomorphismus sein.

Hinweis zur Surjektivität von ρ_3 : Beginne etwa wie folgt: Sei $w_3 \in W_3$ beliebig. Da ρ_4 surjektiv ist, existiert $v_4 \in V_4$ mit $\rho_4(v_4) = \psi_3(w_3)$. Wegen der Kommutativität des rechten Quadrats, folgt $\rho_5(\varphi_4(v_4)) = \psi_4(\rho_4(v_4)) = \psi_4(\psi_3(w_3)) = 0$, denn aufgrund der Exaktheit bei W_4 gilt $\psi_4 \circ \psi_3 = 0$. Da ρ_5 injektiv ist, erhalten wir $\varphi_4(v_4) = 0$, also $v_4 \in \ker(\varphi_4)$. Wegen der Exaktheit bei V_4 existiert $v_3 \in V_3$ mit $\varphi_3(v_3) = v_4$. Wegen der Kommutativität des Diagramms folgt $\psi_3(w_3 - \rho_3(v_3)) = 0$, d.h. $w_3 - \rho_3(v_3) \in \ker(\psi_3)$. Verwende nun die Exaktheit bei W_3 und die Surjektivität von ρ_2 um ein Element $\tilde{v}_3 \in V_3$ mit $\rho_3(\tilde{v}_3) = w_3$ zu konstruieren.

Hinweis zur Injektivität von ρ_3 : Zeige, dass ρ_3 trivialen Kern hat. Sei dazu $v_3 \in V_3$ so, dass $\rho_3(v_3) = 0$. Da ρ_4 injektiv ist, lässt sich daraus $\varphi_3(v_3) = 0$ folgern, also $v_3 \in \ker(\varphi_3)$. Verwende nun die Exaktheit bei V_3 ... dann die Exaktheit bei W_2 ... die Surjektivität von ρ_1 ... und schließlich die Exaktheit bei V_2 .

30. Sofern diese definiert sind, berechne die Produkte $AA, AB, AC, BA, BB, BC, CA, CB$ und CC folgender Matrizen:

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}, \quad C = \begin{pmatrix} 1 & 0 \\ 0 & -1 \\ 2 & 1/2 \end{pmatrix}.$$

Bestimme auch $A^4 = AAAA$.

31. Zeige, dass die folgenden Matrizen über den Körpern \mathbb{Q} , \mathbb{R} und \mathbb{C} invertierbar sind und bestimme ihre Inversen:

$$A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 4 \end{pmatrix}, \quad C = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 & 0 \\ 0 & 0 & 3 & 0 & 0 \\ 0 & 0 & 0 & 4 & 0 \\ 0 & 0 & 0 & 0 & 5 \end{pmatrix}.$$

Welche dieser Matrizen sind über den Körpern \mathbb{Z}_3 , \mathbb{Z}_5 und \mathbb{Z}_7 invertierbar?

32. Sei \mathbb{K} ein Körper und $n \in \mathbb{N}$. Unter einer *Diagonalmatrix* verstehen wir eine Matrix $D \in M_{n \times n}(\mathbb{K})$ der Gestalt

$$D = \begin{pmatrix} a_{11} & 0 & 0 & \cdots & 0 \\ 0 & a_{22} & 0 & \cdots & 0 \\ 0 & 0 & a_{33} & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & 0 \\ 0 & 0 & \cdots & 0 & a_{nn} \end{pmatrix}$$

d.h. $a_{ij} = 0$ falls $i \neq j$. Zeige, dass die Diagonalmatrizen einen Teilraum von $M_{n \times n}(\mathbb{K})$ bilden, der auch abgeschlossen unter Matrizenmultiplikation ist. Zeige, dass für zwei Diagonalmatrizen D und D' stets $DD' = D'D$ gilt. Formuliere ein hinreichendes und notwendiges Kriterium für die Invertierbarkeit von Diagonalmatrizen. Wie sieht im invertierbaren Fall die Inverse aus?

33. Zeige, dass die folgenden beiden Matrizen (über jedem Körper) invertierbar sind und berechne ihre Inversen:

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{pmatrix} \quad \text{und} \quad B = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 0 & 1 & 2 & 3 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

34. Sei \mathbb{K} ein Körper und $n \in \mathbb{N}$. Unter einer *oberen Dreiecksmatrix* verstehen wir eine Matrix $A \in M_{n \times n}(\mathbb{K})$, die folgende Gestalt hat

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ 0 & a_{22} & a_{23} & \cdots & a_{2n} \\ 0 & 0 & a_{33} & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & a_{n-1n} \\ 0 & 0 & \cdots & 0 & a_{nn} \end{pmatrix}$$

d.h. $a_{ij} = 0$ für alle $1 \leq j < i \leq n$. Zeige, dass die oberen Dreiecksmatrizen einen Teilraum von $M_{n \times n}(\mathbb{K})$ bilden, der auch abgeschlossen unter Matrizenmultiplikation ist. Zeige weiters, dass eine obere Dreiecksmatrix genau dann invertierbar ist, wenn alle Diagonalelemente a_{ii} , $i = 1, \dots, n$, verschieden von 0 sind und, dass

in diesem Fall die Inverse Matrix wieder eine obere Dreiecksmatrix bildet. Was kann über die Diagonalelemente der Inversen ausgesagt werden?

35. Betrachte die Matrix $A \in M_{n \times n}(\mathbb{K})$,

$$A = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ & 0 & 1 & \ddots & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ & & & 0 & 1 \\ 0 & \cdots & & & 0 \end{pmatrix}, \quad \text{d.h.} \quad A_{ij} = \begin{cases} 1 & \text{falls } j = i + 1, \text{ und} \\ 0 & \text{andernfalls.} \end{cases}$$

Berechne $A^k = A \cdots A$, für jedes $k \in \mathbb{N}$. *Hinweis: Vielleicht ist es hilfreich dies zunächst für kleine n , etwa $A = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$ oder $A = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}$ durchzuführen.*

36. Unter der Spur einer quadratischen Matrix $A \in M_{n \times n}(\mathbb{K})$ verstehen wir den Skalar $\text{tr}(A) := a_{11} + \cdots + a_{nn}$, d.h. die Summe der Diagonaleinträge. Berechne

$$\text{tr} \begin{pmatrix} 1 & 2 \\ 3 & -1 \end{pmatrix} \quad \text{sowie} \quad \text{tr} \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 2 \\ 3 & 2 & 1 \end{pmatrix}.$$

Zeige, dass die Spur eine surjektive lineare Abbildung $\text{tr}: M_{n \times n}(\mathbb{K}) \rightarrow \mathbb{K}$ definiert. Zeige auch $\text{tr}(AB) = \text{tr}(BA)$, $\text{tr}(A^t) = \text{tr}(A)$ und $\text{tr}(I_n) = n$, für je zwei Matrizen $A, B \in M_{n \times n}(\mathbb{K})$.

37. Zeige, dass die Verknüpfung

$$M_{n \times n}(\mathbb{K}) \times M_{n \times n}(\mathbb{K}) \rightarrow M_{n \times n}(\mathbb{K}), \quad (A, B) \mapsto [A, B] := AB - BA,$$

die folgenden Eigenschaften besitzt:

- (a) $[A_1 + A_2, B] = [A_1, B] + [A_2, B]$ und $[\lambda A, B] = \lambda[A, B]$.
- (b) $[A, B_1 + B_2] = [A, B_1] + [A, B_2]$ und $[A, \lambda B] = \lambda[A, B]$.
- (c) $[A, B] = -[B, A]$ und $[A, A] = 0$.
- (d) $[[A, B], C] + [[B, C], A] + [[C, A], B] = 0$. (Jacobi-Identität)
- (e) $[A, B]^t = -[A^t, B^t]$
- (f) $\text{tr}([A, B]C) = \text{tr}(A[B, C])$

Dabei sind $A, A_1, A_2, B, B_1, B_2, C \in M_{n \times n}(\mathbb{K})$ und $\lambda \in \mathbb{K}$. Der Ausdruck $[A, B] = AB - BA$ wird als *Kommutator* der Matrizen A und B bezeichnet. Zeige auch, dass der Kommutator zweier schiefssymmetrischer Matrizen wieder schiefssymmetrisch ist, d.h. aus $A^t = -A$ und $B^t = -B$ folgt stets $[A, B]^t = -[A, B]$.

38. Betrachte die reellen (2×2) -Matrizen,

$$H = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad E = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \quad \text{und} \quad F = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}.$$

Zeige $[H, E] = 2E$, $[H, F] = -2F$ und $[E, F] = H$, wobei $[A, B] = AB - BA$ den Kommutator bezeichnet, vgl. Aufgabe 37. Zeige auch $[[H, H], E] \neq [H, [H, E]]$ und schlieÙe daraus, dass die Verknüpfung $(A, B) \mapsto [A, B]$ nicht assoziativ ist.

39. Seien $n_1, n_2, m_1, m_2, k_1, k_2 \in \mathbb{N}$ und

$$\begin{aligned} A &\in M_{n_1 \times m_1}(\mathbb{K}), & B &\in M_{n_1 \times m_2}(\mathbb{K}), \\ C &\in M_{n_2 \times m_1}(\mathbb{K}), & D &\in M_{n_2 \times m_2}(\mathbb{K}), \end{aligned}$$

$$\begin{aligned} E &\in M_{m_1 \times k_1}(\mathbb{K}), & F &\in M_{m_1 \times k_2}(\mathbb{K}), \\ G &\in M_{m_2 \times k_1}(\mathbb{K}), & H &\in M_{m_2 \times k_2}(\mathbb{K}). \end{aligned}$$

Setze $n := n_1 + n_2$, $m := m_1 + m_2$, $k := k_1 + k_2$ und betrachte die Blockmatrizen

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix} \in M_{n \times m}(\mathbb{K}) \quad \text{und} \quad \begin{pmatrix} E & F \\ G & H \end{pmatrix} \in M_{m \times k}(\mathbb{K}).$$

Zeige folgende Formel für das Matrizenprodukt:

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix} \begin{pmatrix} E & F \\ G & H \end{pmatrix} = \begin{pmatrix} AE + BG & AF + BH \\ CE + DG & CF + DH \end{pmatrix}$$

Berechne damit

$$\begin{pmatrix} 0 & I_r \\ -I_r & 0 \end{pmatrix}^2 = \begin{pmatrix} 0 & I_r \\ -I_r & 0 \end{pmatrix} \begin{pmatrix} 0 & I_r \\ -I_r & 0 \end{pmatrix},$$

wobei $I_r \in M_{r \times r}(\mathbb{K})$ die $(r \times r)$ -Einheitsmatrix bezeichnet.

40. Sei \mathbb{K} ein Körper, $n_1, n_2 \in \mathbb{N}$ und $n := n_1 + n_2$. Zeige, dass

$$\left\{ \begin{pmatrix} A & B \\ 0 & D \end{pmatrix} \in M_{n \times n}(\mathbb{K}) \mid A \in M_{n_1 \times n_1}(\mathbb{K}), D \in M_{n_2 \times n_2}(\mathbb{K}), B \in M_{n_1 \times n_2}(\mathbb{K}) \right\}$$

einen Teilraum von $M_{n \times n}(\mathbb{K})$ bildet, der auch abgeschlossen unter Matrizenmultiplikation ist. Zeige, dass diese Matrizen genau den linearen Abbildungen $\psi: \mathbb{K}^n \rightarrow \mathbb{K}^n$ entsprechen, für die $\psi(\mathbb{K}^{n_1}) \subseteq \mathbb{K}^{n_1}$ gilt, wobei wir \mathbb{K}^{n_1} in offensichtlicher Weise als Teilraum von \mathbb{K}^n auffassen. Zeige, auch dass eine Matrix dieser Form invertierbar ist, falls A und D beide invertierbar sind und, dass in diesem Fall

$$\begin{pmatrix} A & B \\ 0 & D \end{pmatrix}^{-1} = \begin{pmatrix} A^{-1} & -A^{-1}BD^{-1} \\ 0 & D^{-1} \end{pmatrix}$$

gilt. SchlieÙe daraus, dass

$$\left\{ \begin{pmatrix} A & B \\ 0 & D \end{pmatrix} \in M_{n \times n}(\mathbb{K}) \mid A \in \text{GL}_{n_1}(\mathbb{K}), D \in \text{GL}_{n_2}(\mathbb{K}), B \in M_{n_1 \times n_2}(\mathbb{K}) \right\}$$

eine Untergruppe von $\text{GL}_n(\mathbb{K})$ bildet, vgl. Aufgabe 28. *Hinweis: Aufgabe 39.*

41. Zeige, dass die folgende Matrix über jedem Körper invertierbar ist und bestimme ihre Inverse

$$A = \begin{pmatrix} 2 & 1 & 6 & 7 \\ 3 & 2 & 8 & 4 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 2 & 5 \end{pmatrix}$$

Hinweis: Verwende Aufgabe 40 und Beispiel II.4.8 aus der Vorlesung.

42. Ist $p \in \mathbb{K}[z]$ ein Polynom, d.h. $p = p_0 + p_1z + p_2z^2 + \dots$ mit Koeffizienten $p_i \in \mathbb{K}$, und ist $A \in M_{n \times n}(\mathbb{K})$ eine quadratische Matrix, dann können wir A in p einsetzen und erhalten eine Matrix $p(A) \in M_{n \times n}(\mathbb{K})$,

$$p(A) := p_0I_n + p_1A + p_2A^2 + p_3A^3 + \dots$$

Für beliebige $p, q \in \mathbb{K}[z]$, $\lambda \in \mathbb{K}$ und $A \in M_{n \times n}(\mathbb{K})$ zeige

$$(p + q)(A) = p(A) + q(A), \quad (\lambda p)(A) = \lambda p(A) \quad \text{sowie} \quad (pq)(A) = p(A)q(A).$$

Was bedeutet dies für die Zuordnung $\mathbb{K}[z] \rightarrow F(M_{n \times n}(\mathbb{K}), M_{n \times n}(\mathbb{K}))$, die einem Polynom $p \in \mathbb{K}[z]$ die Abbildung $M_{n \times n}(\mathbb{K}) \rightarrow M_{n \times n}(\mathbb{K})$, $A \mapsto p(A)$, zuordnet? Berechne auch $p(A)$, wobei $p = 2 - 3z + z^2$ und $A = \begin{pmatrix} 1 & 2 \\ 2 & 5 \end{pmatrix}$.

43. Für Vektoren $x, y \in \mathbb{K}^3$ wird ihr Kreuzprodukt $x \times y \in \mathbb{K}^3$ durch

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \times \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} := \begin{pmatrix} x_2y_3 - x_3y_2 \\ x_3y_1 - x_1y_3 \\ x_1y_2 - x_2y_1 \end{pmatrix}$$

definiert. Zeige, dass das Kreuzprodukt folgenden Rechenregeln genügt:

- (a) $x \times (y + \tilde{y}) = x \times y + x \times \tilde{y}$ und $x \times (\lambda y) = \lambda(x \times y)$.
- (b) $(x + \tilde{x}) \times y = x \times y + \tilde{x} \times y$ und $(\lambda x) \times y = \lambda(x \times y)$.
- (c) $x \times y = -y \times x$ und $x \times x = 0$.
- (d) $(x \times y) \times z + (y \times z) \times x + (z \times x) \times y = 0$. (Jacobi-Identität)
- (e) $(x \times y)^t z = x^t(y \times z)$.
- (f) $(x \times y) \times z = (z^t x)y - (y^t z)x$. (Graßmann-Identität)
- (g) $x^t(x \times y) = 0 = y^t(x \times y)$.

Dabei sind $x, \tilde{x}, y, \tilde{y}, z \in \mathbb{K}^3$ und $\lambda \in \mathbb{K}$.

44. Sei \mathbb{K} ein Körper. Zeige, dass die Abbildung

$$\varphi: \mathbb{K}^3 \rightarrow M_{3 \times 3}(\mathbb{K}), \quad \varphi \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} := \begin{pmatrix} 0 & x_1 & x_3 \\ -x_1 & 0 & x_2 \\ -x_3 & -x_2 & 0 \end{pmatrix}$$

einen linearen Isomorphismus auf den Teilraum der schiefsymmetrischen Matrizen definiert. Für beliebige $x, y \in \mathbb{K}^3$ zeige weiters

$$\varphi(x \times y) = [\varphi(x), \varphi(y)].$$

Dabei bezeichnet $[\varphi(x), \varphi(y)] = \varphi(x)\varphi(y) - \varphi(y)\varphi(x)$ den Kommutator von Matrizen, vgl. Aufgabe 37.

45. Sei \mathbb{K} ein Körper. Zeige $W + W' = \mathbb{K}^4$, wobei:

$$W := \{x \in \mathbb{K}^4 \mid x_1 + 2x_2 + 3x_3 + 4x_4 = 0\}$$

$$W' := \{x \in \mathbb{K}^4 \mid x_1 + x_2 + x_3 + x_4 = 0\}$$

46. Seien W, W_1, W_2 und W_3 vier Teilräume eines Vektorraums V . Zeige

- (a) $W_1 + W_2 = W_2 + W_1$
- (b) $W_1 + (W_2 + W_3) = (W_1 + W_2) + W_3$
- (c) $W + \{0\} = W$
- (d) $W_1 + W_2 = W_2 \Leftrightarrow W_1 \subseteq W_2$

Warum bildet die Menge aller Teilräume mit dieser Verknüpfung $+$ i.A. keine Gruppe? Gib einen Vektorraum V an, für den dies doch der Fall ist.

47. Sei \mathbb{K} ein Körper und betrachte folgende Teilräume von \mathbb{K}^5 ,

$$W := \{x \in \mathbb{K}^5 \mid x_4 = x_5 = 0\} \quad \text{und} \quad W' := \{x \in \mathbb{K}^5 \mid x_1 = x_2 = x_3 = 0\},$$

wobei $x_i \in \mathbb{K}$ die Komponenten des Vektors $x \in \mathbb{K}^5$ bezeichnen. Zeige

$$\mathbb{K}^5 = W \oplus W'.$$

Bestimme die Matrizen zur Projektion auf W längs W' , zur Projektion auf W' längs W , zur Spiegelung an W längs W' und zur Spiegelung an W' längs W .

48. Es sei V ein \mathbb{K} -Vektorraum, $\varepsilon: V \rightarrow \mathbb{K}$ linear und $g \in V$, sodass $\varepsilon(g) \neq 0$. Zeige, dass V innere direkte Summe der Teilräume $E := \{v \in V : \varepsilon(v) = 0\}$ und $G := \{\lambda g : \lambda \in \mathbb{K}\}$ ist. Gib auch Formeln für die Projektion auf E längs G , die Projektion auf G längs E und die Spiegelung an E längs G an. *Hinweis: Dies ist eine Verallgemeinerung von Beispiel II.5.10 aus der Vorlesung.*

49. Zeige, dass die Menge der spurfreien Matrizen, d.h.

$$W := \{A \in M_{n \times n}(\mathbb{K}) : \text{tr}(A) = 0\},$$

einen Teilraum von $M_{n \times n}(\mathbb{K})$ bildet. Zeige auch, dass die Vielfachen der Einheitsmatrix,

$$W' := \{\lambda I_n : \lambda \in \mathbb{K}\},$$

einen Teilraum von $M_{n \times n}(\mathbb{K})$ bilden. Unter der Annahme $n \neq 0 \in \mathbb{K}$ zeige weiters

$$M_{n \times n}(\mathbb{K}) = W \oplus W',$$

und gib Formeln für die Projektion auf W längs W' sowie die Projektion auf W' längs W an.

50. Es sei V ein Vektorraum über einem Körper \mathbb{K} in dem $2 \neq 0$ gilt. Weiters sei $\sigma: V \rightarrow V$ linear, sodass $\sigma \circ \sigma = \text{id}_V$. Zeige, dass V innere direkte Summe der beiden Teilräume $W_+ = \{v \in V : \sigma(v) = v\}$ und $W_- = \{v \in V : \sigma(v) = -v\}$ ist. Zeige auch, dass die Projektion auf W_+ längs W_- und die Projektion auf W_- längs W_+ durch $\pi_+ = \frac{1}{2}(\text{id}_V + \sigma)$ bzw. $\pi_- = \frac{1}{2}(\text{id}_V - \sigma)$ gegeben sind. Inwiefern ist diese eine Verallgemeinerung von Beispiel II.5.12 aus der Vorlesung?

51. Betrachte die beiden Teilräume

$$W_1 := \left\{ \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2 \mid x + y = 0 \right\} \quad \text{und} \quad W_2 := \left\{ \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^3 \mid x - 2y = 0 \right\}.$$

Zeige $W_1 \oplus W_2 = \mathbb{R}^2$ und bestimme die Matrix der Projektion auf W_1 längs W_2 sowie die Matrix der Projektion auf W_2 längs W_1 . Bestimme auch die Matrix der Spiegelung an W_1 längs W_2 sowie die Matrix der Spiegelung an W_2 längs W_1 .

52. Betrachte die beiden Teilräume

$$W_1 := \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3 \mid \begin{array}{l} x + y = 0 \\ y + z = 0 \end{array} \right\} \quad \text{und} \quad W_2 := \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3 \mid x + z = 0 \right\}.$$

Zeige $W_1 \oplus W_2 = \mathbb{R}^3$ und bestimme die Matrix der Projektion auf W_1 längs W_2 sowie die Matrix der Projektion auf W_2 längs W_1 . Bestimme auch die Matrix der Spiegelung an W_1 längs W_2 sowie die Matrix der Spiegelung an W_2 längs W_1 .

53. Für einen Teilraum W eines Vektorraums V zeige dass $V/W = \{0\}$ genau dann gilt, wenn $V = W$.

54. Sei $W := \{x \in \mathbb{R}^3 \mid x_1 + x_2 + x_3 = 0\}$ und bezeichne $\pi: \mathbb{R}^3 \rightarrow \mathbb{R}^3/W$ die kanonische Projektion. Zeige $\pi(u) = \pi(v)$ und $\pi(u) \neq \pi(w)$, wobei:

$$u = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}, \quad v := \begin{pmatrix} 3 \\ 4 \\ -1 \end{pmatrix}, \quad w = \begin{pmatrix} 1 \\ 2 \\ 4 \end{pmatrix}.$$

55. Sei W ein Teilraum eines Vektorraums V und $\varphi: V \rightarrow V$ linear, sodass $\varphi(W) \subseteq W$. Zeige, dass es genau eine lineare Abbildung $\bar{\varphi}: V/W \rightarrow V/W$ gibt, sodass $\pi \circ \varphi = \bar{\varphi} \circ \pi$, wobei $\pi: V \rightarrow V/W$ die kanonische Projektion bezeichnet.

56. Sind A und B zwei Teilmengen eines Vektorraums V , dann gilt für die lineare Hülle i.A. $\langle A \cap B \rangle \neq \langle A \rangle \cap \langle B \rangle$. Erläutere dies an einem Beispiel.

57. Welche der folgenden Teilmengen bilden Erzeugendensysteme der angegebenen Vektorräume?

$$\begin{array}{ll} \left\{ \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 16 \\ 17 \end{pmatrix} \right\} \subseteq \mathbb{R}^2 & \left\{ \begin{pmatrix} 2 \\ 5 \end{pmatrix}, \begin{pmatrix} 1 \\ 3 \end{pmatrix} \right\} \subseteq \mathbb{R}^2 \\ \left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 2 \end{pmatrix}, \begin{pmatrix} 7 \\ 8 \\ 9 \end{pmatrix} \right\} \subseteq \mathbb{R}^3 & \left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}, \begin{pmatrix} 1 \\ 4 \\ 5 \end{pmatrix} \right\} \subseteq \mathbb{R}^3 \\ \left\{ \begin{pmatrix} 1 \\ x \end{pmatrix} : x \in \mathbb{R} \right\} \subseteq \mathbb{R}^2 & \left\{ \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} i \\ 0 \end{pmatrix} \right\} \subseteq \mathbb{C}^2 \\ \left\{ \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 2 \\ 2 \\ 1 \end{pmatrix}, \begin{pmatrix} 3 \\ 3 \\ 1 \end{pmatrix} \right\} \subseteq \mathbb{R}^3 & \left\{ \begin{pmatrix} 1 \\ i \end{pmatrix}, \begin{pmatrix} i \\ -1 \end{pmatrix}, \begin{pmatrix} -1 \\ -i \end{pmatrix}, \begin{pmatrix} -i \\ 1 \end{pmatrix} \right\} \subseteq \mathbb{C}^2 \end{array}$$

58. Zeige, dass die folgenden Teilmengen linear abhängig sind und gib jeweils ein Element an, das sich als Linearkombination der restlichen schreiben lässt:

$$\begin{array}{ll} \left\{ \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} -1 \\ -1 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \end{pmatrix} \right\} \subseteq \mathbb{R}^2 & \left\{ \begin{pmatrix} 2 \\ -1 \\ 5 \end{pmatrix}, \begin{pmatrix} 3 \\ 1 \\ 4 \end{pmatrix}, \begin{pmatrix} 1 \\ -3 \\ 6 \end{pmatrix} \right\} \subseteq \mathbb{R}^3 \\ \left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 3 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right\} \subseteq \mathbb{R}^4 & \left\{ \begin{pmatrix} 7 \\ i \\ 4-i \end{pmatrix}, \begin{pmatrix} 7+7i \\ i-1 \\ 5+3i \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 2i \end{pmatrix} \right\} \subseteq \mathbb{C}^3 \end{array}$$

59. Zeige, dass die folgenden Teilmengen linear unabhängig sind:

$$\left\{ \begin{pmatrix} 9 \\ 7 \end{pmatrix}, \begin{pmatrix} 6 \\ 5 \end{pmatrix} \right\} \subseteq \mathbb{R}^2 \quad \left\{ \begin{pmatrix} -1 \\ 3 \\ 5 \end{pmatrix}, \begin{pmatrix} 0 \\ 2 \\ 0 \end{pmatrix}, \begin{pmatrix} 2 \\ 4 \\ -3 \end{pmatrix} \right\} \subseteq \mathbb{R}^3$$

$$\left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \\ 1 \end{pmatrix} \right\} \subseteq \mathbb{R}^4 \quad \left\{ \begin{pmatrix} i \\ 0 \\ 1+i \end{pmatrix}, \begin{pmatrix} 2+7i \\ 3-i \\ 5 \end{pmatrix} \right\} \subseteq \mathbb{C}^3$$

60. Fasse $V = \mathbb{R}$ als Vektorraum über \mathbb{Q} auf und zeige, dass $\{1, \sqrt{2}\} \subseteq V$ eine linear unabhängige Teilmenge bildet.

61. Welche der folgenden Teilmengen bilden Basen der angegebenen Vektorräume?

$$\left\{ \begin{pmatrix} 1 \\ 7 \end{pmatrix}, \begin{pmatrix} 2 \\ 13 \end{pmatrix} \right\} \subseteq \mathbb{R}^2 \quad \left\{ \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}, \begin{pmatrix} 2 \\ 3 \\ 4 \end{pmatrix} \right\} \subseteq \mathbb{R}^3$$

$$\left\{ \begin{pmatrix} 1 \\ i \end{pmatrix}, \begin{pmatrix} i \\ i \end{pmatrix} \right\} \subseteq \mathbb{C}^2 \quad \{1, 1+z, 1+2z+z^2\} \subseteq \mathbb{R}[z]_{\leq 2}$$

62. Erweitere den Vektor $\begin{pmatrix} 5 \\ 3 \\ 1 \end{pmatrix}$ zu einer Basis von \mathbb{R}^3 .

63. Zeige, dass jede endliche totalgeordnete Menge ein eindeutiges Maximum besitzt. *Hinweis: Induktion nach der Anzahl der Elemente.*

64. Sei W ein Teilraum eines Vektorraums V . Zeige, dass ein zu W komplementärer Teilraum M in V existiert, d.h. $W \oplus M = V$, wie folgt:

- (a) Sei U ein Teilraum von V , $W \cap U = \{0\}$ und $v \in V \setminus (W + U)$. Zeige, dass $U' := \langle v \rangle + U$ ein Teilraum von V ist, für den $U' \supsetneq U$ und $W \cap U' = \{0\}$ gilt.
- (b) Die Mengeninklusion \subseteq definiert eine Halbordnung auf der Menge

$$X = \{U \mid U \text{ ist Teilraum von } V \text{ und } W \cap U = \{0\}\}.$$

Zeige, dass jede Kette in X eine obere Schranke besitzt. *Hinweis: Für jede totalgeordnete Teilmenge $K \subseteq X$ ist $S := \bigcup_{U \in K} U$ ein Teilraum von V für den $W \cap S = \{0\}$ gilt, vgl. Aufgabe 20.*

- (c) Nach dem Lemma von Zorn existiert ein maximales Element $M \in X$. Zeige mit Hilfe von (a), dass $W \oplus M = V$ gilt.

65. Seien $U \xrightarrow{\varphi} V \xrightarrow{\psi} W$ linear und $U^* \xleftarrow{\varphi^t} V^* \xleftarrow{\psi^t} W^*$ die dazu dualen Abbildungen. Zeige

$$\text{img}(\varphi) = \ker(\psi) \Leftrightarrow \text{img}(\varphi^t) = \ker(\psi^t).$$

Hinweis: Verwende Satz III.4.10 aus der Vorlesung.

66. a) Sei $\pi: V \rightarrow V$ ein Projektor, d.h. $\pi \circ \pi = \pi$. Zeige, dass dann auch die duale Abbildung $\pi^t: V^* \rightarrow V^*$ ein Projektor ist, und beschreibe dessen Bild und Kern. b) Seien nun W_1 und W_2 zwei komplementäre Teilräume von V , d.h. $V = W_1 \oplus W_2$. Weiters bezeichne $\pi_1: V \rightarrow W_1$ die damit assoziierte Projektion auf W_1 längs W_2 und $\pi_2: V \rightarrow W_2$ die Projektion auf W_2 längs W_1 . Nach Satz III.4.9 ist dann auch $V^* = W_1^\circ \oplus W_2^\circ$. Zeige, dass π_1^t mit der Projektion auf W_2° längs W_1° übereinstimmt, und analog für π_2^t .

67. Sei V ein Vektorraum und $n \in \mathbb{N}_0$. Zeige, dass die folgenden Aussagen äquivalent sind:

- (a) $\dim(V) = n$.
- (b) Es existieren n linear unabhängige Vektoren, und je $n + 1$ Vektoren in V sind linear abhängig.
- (c) Es existiert ein Erzeugendensystem von V mit n Vektoren, und je $n - 1$ Vektoren erzeugen V nicht.

68. Sei X eine n -elementige Menge. Zeige $\dim(F(X, \mathbb{K})) = n$.

69. Sei X eine Menge mit unendlich vielen Elementen. Zeige, dass $F(X, \mathbb{K})$ unendlich-dimensional ist. *Hinweis: Für jedes $x \in X$ betrachte die Funktion $e_x: X \rightarrow \mathbb{K}$, $e_x(x) := 1$, $e_x(y) := 0$ falls $y \neq x$, und zeige, dass die Menge $\{e_x \mid x \in X\}$ linear unabhängig in $F(X, \mathbb{K})$ ist.*

70. Zeige: Zwei Vektoren $x, y \in \mathbb{R}^3$ sind genau dann linear unabhängig, wenn ihr Kreuzprodukt $z := x \times y$ verschieden von Null ist. In diesem Fall bilden die Vektoren x, y, z eine Basis von \mathbb{R}^3 . *Hinweis: Es genügt die lineare Unabhängigkeit zu zeigen.* In diesem Fall gilt weiters

$$\left\{ \begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix} \in \mathbb{R}^3 \mid \begin{array}{l} x_1 u_1 + x_2 u_2 + x_3 u_3 = 0 \\ y_1 u_1 + y_2 u_2 + y_3 u_3 = 0 \end{array} \right\} = \{ \lambda z \mid \lambda \in \mathbb{R} \},$$

sowie

$$\{ \lambda x + \mu y \mid \lambda, \mu \in \mathbb{R} \} = \left\{ \begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix} \in \mathbb{R}^3 \mid z_1 u_1 + z_2 u_2 + z_3 u_3 \right\}.$$

Hinweis: In beiden Fällen folgt eine Inklusion sofort aus $x^t z = 0 = y^t z$, siehe Aufgabe 43. Die Gleichheit folgt dann aus Dimensionsgründen.

71. Seien W_1 und W_2 zwei verschiedene $(n - 1)$ -dimensionale Teilräume von \mathbb{K}^n . Zeige $W_1 + W_2 = \mathbb{K}^n$ und $\dim(W_1 \cap W_2) = n - 2$. *Hinweis: Beispiel IV.2.6.*

72. Sei V ein Vektorraum und $\alpha, \beta \in V^*$ zwei nicht-triviale lineare Funktionale, $\alpha \neq 0$, $\beta \neq 0$. Zeige, dass die Hyperebenen $\ker(\alpha)$ und $\ker(\beta)$ genau dann übereinstimmen, wenn $\lambda \in \mathbb{K}$ existiert, sodass $\beta = \lambda \alpha$.

73. Sei $0 \rightarrow W \xrightarrow{\varphi} V \xrightarrow{\psi} U \rightarrow 0$ eine kurze exakte Sequenz linearer Abbildungen, d.h. φ sei injektiv, ψ surjektiv und $\text{img}(\varphi) = \ker(\psi)$. Zeige, dass V genau dann endlich-dimensional ist, wenn W und U beide endlich-dimensional sind, und in diesem Fall gilt

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

Hinweis: Zeige $U \cong V / \text{img}(\varphi)$, $W \cong \text{img}(\varphi)$ und verwende Korollar IV.2.8.

74 (Kodimension und Durchschnitt). Ein Teilraum W eines Vektorraums V hat endliche *Kodimension in V* , falls V/W endlich-dimensional ist. In diesem Fall wird $\text{codim}_V(W) := \dim(V/W)$ die *Kodimension von W in V genannt*. Seien nun W_1 und W_2 zwei Teilräume von V . Zeige, dass $W_1 \cap W_2$ genau dann endliche Kodimension in V hat, wenn W_1 und W_2 beide endliche Kodimension in V haben. Zeige weiters, dass in diesem Fall auch $W_1 + W_2$ endliche Kodimension in V hat und es gilt

$$\text{codim}_V(W_1 \cap W_2) + \text{codim}_V(W_1 + W_2) = \text{codim}_V(W_1) + \text{codim}_V(W_2).$$

Im transversalen Fall, d.h. wenn $W_1 + W_2 = V$ gilt, erhalten wir

$$\text{codim}_V(W_1 \cap W_2) = \text{codim}_V(W_1) + \text{codim}_V(W_2).$$

Hinweis: $\text{codim}_V(W) = \dim(V/W) = \dim(W^\circ)$, *Satz IV.2.1 und Satz III.4.9.*

75. Erläutere die Gleichungen

$$\dim_{\mathbb{C}}(\mathbb{C}^n) = n \quad \text{und} \quad \dim_{\mathbb{R}}(\mathbb{C}^n) = 2n.$$

Hinweis: $\mathbb{C}^n = \mathbb{R}^{2n}$, *siehe auch Beispiel IV.1.12.*

76. Sei V ein endlich-dimensionaler komplexer Vektorraum und b_1, \dots, b_n eine Basis von V . Zeige, dass $b_1, \mathbf{i}b_1, b_2, \mathbf{i}b_2, \dots, b_n, \mathbf{i}b_n$ eine Basis des V zugrundeliegenden reellen Vektorraums $V^{\mathbb{R}}$ bildet, vgl. Bemerkung IV.2.21, und schlieÙe daraus $\dim_{\mathbb{R}}(V^{\mathbb{R}}) = 2 \dim_{\mathbb{C}}(V)$.

77. Führe die Details in Bemerkung IV.2.22 aus.

78. Es sei V ein reeller Vektorraum und $J: V \rightarrow V$ linear mit $J \circ J = -\text{id}_V$. Zeige, dass V bezüglich der Skalarmultiplikation

$$\mathbb{C} \times V \rightarrow V, \quad (a + \mathbf{b}i)v := av + bJ(v),$$

zu einem komplexen Vektorraum wird. SchlieÙe daraus, dass dies nur möglich ist, wenn $\dim_{\mathbb{R}}(V)$ gerade ist. *Hinweis:* Für den letzten Teil verwende Aufgabe 76.

79. Sei

$$\{0\} \xrightarrow{\varphi_{-1}=0} V_0 \xrightarrow{\varphi_0} V_1 \xrightarrow{\varphi_1} V_2 \xrightarrow{\varphi_2} \dots \rightarrow V_{n-1} \xrightarrow{\varphi_{n-1}} V_n \xrightarrow{\varphi_n=0} \{0\}$$

eine Folge linearer Abbildungen zwischen endlich-dimensionalen Vektorräumen, sodass $\varphi_i \circ \varphi_{i-1} = 0$, für jedes i . SchlieÙe daraus $\text{img}(\varphi_{i-1}) \subseteq \ker(\varphi_i)$ und definiere Vektorräume $H_i := \ker(\varphi_i) / \text{img}(\varphi_{i-1})$. Zeige nun

$$\sum_i (-1)^i \dim(V_i) = \sum_i (-1)^i \dim(H_i).$$

Hinweis: Verwende $\text{img}(\varphi_i) \cong V_i / \ker(\varphi_i)$ und *Korollar IV.2.8.*

80. Seien $A \in M_{m \times n}(\mathbb{K})$, $S \in \text{GL}_n(\mathbb{K})$ und $T \in \text{GL}_m(\mathbb{K})$. Zeige

$$\text{rank}(TAS) = \text{rank}(A).$$

Hinweis: Betrachte die mit den Matrizen assoziierten linearen Abbildungen und verwende Proposition IV.2.23.

81. Sei W ein Teilraum eines endlich-dimensionalen Vektorraums V , und bezeichne $\iota: V \rightarrow V^{**}$, $\iota(v)(\alpha) = \alpha(v)$, die natürliche Abbildung, $v \in V$, $\alpha \in V^*$. Zeige $\iota(W) = W^{\circ\circ}$. *Hinweis: Proposition III.4.13.*

82. Seien $\varphi: V \rightarrow W$ und $\psi: W \rightarrow U$ zwei lineare Abbildungen mit endlichem Rang. Zeige, dass dann auch $\psi \circ \varphi: V \rightarrow U$ endlichen Rang hat und

$$\text{rank}(\psi \circ \varphi) \leq \min\{\text{rank}(\psi), \text{rank}(\varphi)\}$$

gelten muss. Schließe daraus, dass für je zwei Matrizen $A \in M_{m \times n}(\mathbb{K})$ und $B \in M_{n \times l}(\mathbb{K})$ folgende Ungleichung gilt:

$$\text{rank}(AB) \leq \min\{\text{rank}(A), \text{rank}(B)\}$$

83. Bestimme die Dimension sowie eine Basis des von den Vektoren

$$v_1 = \begin{pmatrix} 0 \\ 1 \\ 2 \\ -1 \\ -1 \\ 2 \end{pmatrix}, \quad v_2 = \begin{pmatrix} 0 \\ 1 \\ 2 \\ 1 \\ 7 \\ 4 \end{pmatrix}, \quad v_3 = \begin{pmatrix} 0 \\ 2 \\ 4 \\ 0 \\ 6 \\ 6 \end{pmatrix}, \quad v_4 = \begin{pmatrix} 0 \\ 3 \\ 6 \\ -5 \\ -11 \\ 4 \end{pmatrix}, \quad v_5 = \begin{pmatrix} 0 \\ 4 \\ 8 \\ 0 \\ 12 \\ 15 \end{pmatrix}$$

aufgespannten Teilraums $W = \langle v_1, v_2, v_3, v_4, v_5 \rangle \subseteq \mathbb{R}^6$. Gib ein minimales Gleichungssystem für W an. Bestimme eine Basis von W , die aus gewissen der Vektoren v_i besteht. Gib schließlich auch ein Komplement von W an und beschreibe dieses mit einer Basis und durch ein Gleichungssystem. *Hinweis: siehe Beispiel IV.3.17.*

84. Zeige, dass die Vektoren

$$v_1 = \begin{pmatrix} 1 \\ 1 \\ -1 \\ 2 \end{pmatrix}, \quad v_2 = \begin{pmatrix} 2 \\ 3 \\ -1 \\ 6 \end{pmatrix}, \quad v_3 = \begin{pmatrix} 3 \\ 3 \\ -3 \\ 6 \end{pmatrix}, \quad v_4 = \begin{pmatrix} 4 \\ 4 \\ -2 \\ 8 \end{pmatrix}, \quad v_5 = \begin{pmatrix} 6 \\ 11 \\ 0 \\ 24 \end{pmatrix}$$

ein Erzeugendensystem von \mathbb{R}^4 bilden, und bestimme vier dieser Vektoren v_i , die eine Basis von \mathbb{R}^4 bilden. *Hinweis: siehe Beispiel IV.3.22.*

85. Zeige, dass die Vektoren

$$v_1 = \begin{pmatrix} 1 \\ -1 \\ 1 \\ -1 \\ 1 \end{pmatrix}, \quad v_2 = \begin{pmatrix} 1 \\ 2 \\ 1 \\ 2 \\ 1 \end{pmatrix}, \quad v_3 = \begin{pmatrix} 1 \\ 2 \\ 3 \\ 1 \\ 2 \end{pmatrix}$$

linear unabhängig in \mathbb{R}^5 sind und erweitere sie zu einer Basis von \mathbb{R}^5 . *Hinweis: Siehe Beispiel IV.3.18.*

86. Es bezeichne W den Teilraum aller $y \in \mathbb{R}^6$, für die das Gleichungssystem

$$\begin{aligned} 2x_1 + x_2 + 3x_3 + 6x_4 + 10x_5 &= y_1 \\ 3x_1 + x_2 + x_3 + 5x_4 + 7x_5 &= y_2 \\ 5x_1 + 2x_2 + 4x_3 + 11x_4 + 17x_5 &= y_3 \\ 7x_1 + 3x_2 + x_3 + 11x_4 + 15x_5 &= y_4 \\ 11x_1 + 5x_2 + 5x_3 + 21x_4 + 31x_5 &= y_5 \\ 13x_1 + 8x_2 + 9x_3 + 30x_4 + 47x_5 &= y_6 \end{aligned}$$

lösbar ist. Bestimme $\dim(W)$ und eine Basis von W . Gib auch ein minimales Gleichungssystem für W an. *Hinweis: Siehe Beispiel IV.3.23.*

87. Bezeichne $L \subseteq \mathbb{R}^6$ den Lösungsraum des Gleichungssystems:

$$\begin{aligned} x_1 - x_2 + x_3 + 2x_4 - 2x_5 + 2x_6 &= 0 \\ 2x_1 - 2x_2 + 2x_3 + 4x_4 - 4x_5 + 4x_6 &= 0 \\ 2x_1 - 2x_2 + 2x_3 + 3x_4 - 3x_5 + 3x_6 &= 0 \\ x_1 - x_2 + x_3 + 5x_4 - 5x_5 + 5x_6 &= 0 \end{aligned}$$

Bestimme $\dim(L)$ und eine Basis von L . Gib auch ein minimales Gleichungssystem für L an. Bestimme auch ein minimales Gleichungssystem für L , das aus einigen der ursprünglichen Gleichungen besteht. Bestimme ein Komplement von L und beschreibe es mit Hilfe einer Basis als auch durch ein Gleichungssystem. *Hinweis: Siehe Beispiel IV.3.19.*

88. Bestimme alle Lösungen des Gleichungssystems

$$\begin{aligned} 3x_1 - 3x_2 - 3x_3 - 12x_4 - 12x_5 &= -21 \\ x_1 - 3x_2 - 5x_3 - 18x_4 - 20x_5 &= -41 \\ 2x_1 - x_2 + 2x_3 + 3x_4 + 6x_5 &= 15 \\ -x_1 + 3x_2 + 4x_3 + 16x_4 + 17x_5 &= 35 \end{aligned}$$

und gib diese in Parameterform an. Bestimme auch ein minimales Gleichungssystem für den Lösungsraum. Welche Dimension hat dieser? *Hinweis: Siehe Beispiel IV.4.2.*

89. Bestimme alle Lösungen des Gleichungssystems:

$$\begin{aligned} x_1 + x_2 + 3x_3 + 3x_4 + 5x_5 &= 37 \\ x_1 + 2x_2 + 5x_3 + 6x_4 + 8x_5 &= 60 \\ 2x_1 + x_2 + 4x_3 + 3x_4 + 8x_5 &= 59 \\ -x_1 - 3x_2 - 7x_3 - 9x_4 - 8x_5 &= -59 \end{aligned}$$

Gib auch eine Basis des Lösungsraums für das assoziierte homogene System an.

90. Bestimme die Dimension des affinen Teilraums

$$E = \left(\begin{pmatrix} 1 \\ -1 \\ 1 \\ -1 \\ 1 \end{pmatrix} + \left\langle \left(\begin{pmatrix} 1 \\ 2 \\ 3 \\ 4 \\ 5 \end{pmatrix}, \begin{pmatrix} 5 \\ 4 \\ 3 \\ 2 \\ 1 \end{pmatrix}, \begin{pmatrix} 7 \\ 8 \\ 9 \\ 10 \\ 11 \end{pmatrix}, \begin{pmatrix} 11 \\ 10 \\ 9 \\ 8 \\ 7 \end{pmatrix} \right\rangle \right)$$

von \mathbb{R}^5 . Gib auch ein minimales Gleichungssystem für E an. *Hinweis: Siehe Beispiel IV.4.4.*

91. Bestimme die Inversen folgender reeller Matrizen:

$$A = \begin{pmatrix} 2 & -1 & 1 \\ 1 & 0 & -1 \\ 0 & 1 & -2 \end{pmatrix}, \quad B = \begin{pmatrix} 2 & 0 & 0 & 1 \\ 4 & 3 & 1 & 2 \\ 0 & -3 & 3 & 0 \\ 0 & 3 & -3 & 5 \end{pmatrix}$$

92. Bestimme die Inverse folgender komplexer Matrix:

$$A = \begin{pmatrix} \mathbf{i} & 3 & 2 + \mathbf{i} \\ 2\mathbf{i} & 7 & -\mathbf{i} \\ -1 & -1 + 4\mathbf{i} & 7 + \mathbf{i} \end{pmatrix}.$$

93. Bestimme die Inversen folgender Matrizen über dem Körper $\mathbb{K} = \mathbb{Z}_2$:

$$A = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 1 \end{pmatrix}.$$

94. Zeige, dass die Vektoren

$$b_1 = \begin{pmatrix} 2 \\ 2 \\ 2 \\ 9 \end{pmatrix}, \quad b_2 = \begin{pmatrix} 7 \\ 8 \\ 8 \\ 0 \end{pmatrix}, \quad b_3 = \begin{pmatrix} 1 \\ 1 \\ 4 \\ 4 \end{pmatrix}, \quad b_4 = \begin{pmatrix} 8 \\ 8 \\ 5 \\ 5 \end{pmatrix}$$

eine Basis von \mathbb{R}^4 bilden. *Hinweis: Siehe Beispiel IV.5.2.*

95. Zeige, dass das Gleichungssystem

$$\begin{aligned} x_1 + x_2 + x_3 + x_4 &= y_1 \\ x_2 + 2x_3 + 3x_4 &= y_2 \\ x_2 + 4x_3 + 9x_4 &= y_3 \\ x_2 + 8x_3 + 27x_4 &= y_4 \end{aligned}$$

für alle $y_1, y_2, y_3, y_4 \in \mathbb{R}$ eindeutig lösbar ist und gib diese Lösung an. *Hinweis: Siehe Beispiel IV.5.6.*

96. Es sei A eine invertierbare Matrix mit rationalen Einträgen. Erkläre warum dann auch die Inverse, A^{-1} , nur rationale Einträge besitzt. Gib eine invertierbare (2×2) -Matrix mit ganzzahligen Einträgen an, deren Inverse nicht-ganzzahlige Einträge hat.

97. Sei V ein endlich-dimensionaler Vektorraum. Weiters seien B und C zwei geordnete Basen von V und es bezeichnen B^* bzw. C^* die dazu dualen Basen von V^* . Zeige, dass aus $B^* = C^*$ schon $B = C$ folgt. *Hinweis: Betrachte die zu B^* und C^* dualen Basen B^{**} und C^{**} von V^{**} und verwende Lemma IV.6.7 aus der Vorlesung.*

98. Zeige, dass $B = (b_1, b_2, b_3, b_4)$ eine Basis von \mathbb{R}^4 bildet, wobei

$$b_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}, \quad b_2 = \begin{pmatrix} 0 \\ 2 \\ 2 \\ 2 \end{pmatrix}, \quad b_3 = \begin{pmatrix} 0 \\ 0 \\ 3 \\ 3 \end{pmatrix}, \quad b_4 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 4 \end{pmatrix}.$$

Bestimme die Basiswechsellmatrix T_{BE} , wobei $E = (e_1, e_2, e_3, e_4)$ die Standardbasis bezeichnet, und berechne damit die Koordinaten $[v]_B$ folgender Vektoren bezüglich B :

$$v = \begin{pmatrix} 5 \\ 7 \\ 9 \\ 0 \end{pmatrix}, \quad v = \begin{pmatrix} -3 \\ 7 \\ 3 \\ -1 \end{pmatrix}, \quad v = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}, \quad v = \begin{pmatrix} 6 \\ 5 \\ 4 \\ 2 \end{pmatrix}, \quad v = \begin{pmatrix} -1 \\ 1 \\ -1 \\ 1 \end{pmatrix}.$$

99. Zeige, dass $B = (b_1, b_2, b_3)$ eine Basis von \mathbb{R}^3 bildet, wobei

$$b_1 = \begin{pmatrix} 0 \\ 1 \\ 2 \end{pmatrix}, \quad b_2 = \begin{pmatrix} -1 \\ 2 \\ 5 \end{pmatrix}, \quad b_3 = \begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix}.$$

Bestimme die Matrix (bezüglich der Standardbasis) einer linearen Abbildung $\varphi: \mathbb{R}^3 \rightarrow \mathbb{R}^3$, für die $\varphi(b_1) = b_1$, $\varphi(b_2) = 2b_2$ und $\varphi(b_3) = 3b_3$ gilt. Wieviele solche Abbildungen φ gibt es? *Hinweis: Ohne Rechnung lassen sich $[\varphi]_{BB}$ und T_{EB} angeben, wobei $E = (e_1, e_2, e_3)$ die Standardbasis von \mathbb{R}^3 bezeichnet.*

100. Zeige, dass die Vektoren

$$b_1 = \begin{pmatrix} 1 \\ -1 \\ 1 \\ -1 \end{pmatrix}, \quad b_2 = \begin{pmatrix} 0 \\ 1 \\ -1 \\ 1 \end{pmatrix}, \quad b_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \\ -1 \end{pmatrix}, \quad b_4 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}$$

eine Basis von \mathbb{R}^4 bilden. SchlieÙe daraus, dass \mathbb{R}^4 direkte Summe der beiden Teilräume $W = \langle b_1, b_2 \rangle$ und $W' = \langle b_3, b_4 \rangle$ ist, $\mathbb{R}^4 = W \oplus W'$. Bestimme die Matrix (bezüglich der Standardbasis) des damit assoziierten Projektors $\pi: \mathbb{R}^4 \rightarrow \mathbb{R}^4$ auf W längs W' . Berechne damit auch die Matrix des komplementären Projektors, sowie die Matrizen der beiden damit assoziierten Spiegelungen. *Hinweis: Ohne Rechnung lassen sich $[\pi]_{BB}$ und T_{EB} angeben, wobei $E = (e_1, e_2, e_3, e_4)$ die Standardbasis von \mathbb{R}^4 bezeichnet.*

101. Zeige, dass die Funktionen $f_1(x) := \sin(2x)$, $f_2(x) := \cos(2x)$, $f_3(x) := \sin(3x)$ und $f_4(x) := \cos(3x)$ linear unabhängig in $F(\mathbb{R}, \mathbb{R})$ sind und daher eine Basis $B = (f_1, f_2, f_3, f_4)$ des von ihnen aufgespannten Teilraums $V := \langle f_1, f_2, f_3, f_4 \rangle \subseteq F(\mathbb{R}, \mathbb{R})$ bilden. Zeige, dass die Ableitung eine lineare Abbildung $D: V \rightarrow V$, $D(f) := f'$, liefert, und bestimme die Matrix $[D]_{BB}$ von D bezüglich B .

102. Zeige, dass die Polynome

$$\begin{aligned} h_0 &= 1 \\ h_1 &= z \\ h_2 &= z^2 - 1 \\ h_3 &= z^3 - 3z \\ h_4 &= z^4 - 6z^2 + 3 \end{aligned}$$

eine Basis $H = (h_0, h_1, h_2, h_3, h_4)$ von $\mathbb{R}[z]_{\leq 4}$ bilden. Bestimme die Matrix $[L]_{HH}$ der linearen Abbildung

$$L: \mathbb{R}[z]_{\leq 4} \rightarrow \mathbb{R}[z]_{\leq 4}, \quad D(p) := p'' - zp'.$$

Berechne den Rang der linearen Abbildung L und gib eine Basis ihres Kerns an. Bestimme jenes Polynom $p \in \mathbb{R}[z]_{\leq 4}$ für das $[p]_H = (1, 0, 1, 0, 1)^t$ gilt. Berechne auch die Koordinaten $[q]_H$ des Polynoms $q = z^4 + 2z^3 - 6z^2 - 6z - 6$.

103. Betrachte die drei linearen Funktionale $\gamma_1, \gamma_2, \gamma_3: \mathbb{R}^3 \rightarrow \mathbb{R}$,

$$\gamma_1 \begin{pmatrix} x \\ y \\ z \end{pmatrix} = x + 2y + 3z, \quad \gamma_2 \begin{pmatrix} x \\ y \\ z \end{pmatrix} = x + 4y + 9z, \quad \gamma_3 \begin{pmatrix} x \\ y \\ z \end{pmatrix} = x + 8y + 27z.$$

Zeige, dass $C = (\gamma_1, \gamma_2, \gamma_3)$ eine Basis von $(\mathbb{R}^3)^*$ bildet. Bestimme eine Basis $B = (b_1, b_2, b_3)$ von \mathbb{R}^3 , deren duale Basis mit C übereinstimmt, $B^* = C$. *Hinweis: Ohne Rechnung lässt sich die Basiswechselmatrix $T_{E^*C} = T_{E^*B^*}$ angeben, wobei E^* die zur Standardbasis $E = (e_1, e_2, e_3)$ duale Basis von $(\mathbb{R}^3)^*$ bezeichnet. Berechne daraus T_{EB} , siehe Korollar IV.6.21, und lies die Basis B ab.*

104. Es sei $S \in \text{GL}_n(\mathbb{K})$ eine invertierbare Matrix. Zeige, dass geordnete Basen B und C von \mathbb{K}^n existieren, sodass $T_{CB} = S$ gilt, wobei T_{CB} die Basiswechselmatrix bezeichnet. *Hinweis: Wir können für B (oder C) die Standardbasis von \mathbb{K}^n verwenden.*

105 (Äquivalente Matrizen). Zwei Matrizen $A, B \in M_{m \times n}(\mathbb{K})$ werden äquivalent genannt, falls invertierbare Matrizen $S \in \text{GL}_n(\mathbb{K})$ und $T \in \text{GL}_m(\mathbb{K})$ existieren, sodass $B = T A S$. Zeige, dass dies eine Äquivalenzrelation auf der Menge der $(m \times n)$ -Matrizen definiert. Zeige weiters: Ist $\psi: V \rightarrow W$ eine lineare Abbildung zwischen endlich-dimensionalen \mathbb{K} -Vektorräumen, C, \tilde{C} geordnete Basen von V und D, \tilde{D} geordnete Basis von W , dann sind die Matrizen $[\psi]_{DC}$ und $[\psi]_{\tilde{D}\tilde{C}}$ ähnlich. Zeige auch, dass die folgenden Aussagen äquivalent sind:

- (a) A und B sind äquivalent.
- (b) B lässt sich aus A durch Zeilen- und Spaltenumformungen gewinnen.
- (c) Es existieren Basen C von \mathbb{K}^m und D von \mathbb{K}^n , sodass $B = [\psi_A]_{DC}$, wobei ψ_A die lineare Abbildung $\psi_A: \mathbb{K}^n \rightarrow \mathbb{K}^m$, $\psi_A(x) = Ax$, bezeichnet.
- (d) A und B haben gleichen Rang.

Schließe daraus, dass jede Matrix $A \in M_{m \times n}(\mathbb{K})$ äquivalent zu einer Matrix der Form $\begin{pmatrix} I_k & 0 \\ 0 & 0 \end{pmatrix} \in M_{m \times n}(\mathbb{K})$ ist, wobei $k = \text{rank}(A)$.

106 (Ähnliche Matrizen). Zwei quadratische Matrizen $A, B \in M_{n \times n}(\mathbb{K})$ werden ähnlich genannt, falls eine invertierbare Matrix $S \in \text{GL}_n(\mathbb{K})$ existiert, sodass $B = S A S^{-1}$. Zeige, dass dies eine Äquivalenzrelation auf der Menge der $(n \times n)$ -Matrizen definiert. Zeige weiters: Ist $\varphi: V \rightarrow V$ ein Endomorphismus eines endlich-dimensionalen Vektorraums V und sind B bzw. C zwei geordnete Basen von V , dann sind die Matrizen $[\varphi]_{BB}$ und $[\varphi]_{CC}$ ähnlich. Zeige auch, dass für fixes $\lambda \in \mathbb{K}$ die beiden Matrizen $A = \begin{pmatrix} \lambda & 0 \\ 0 & \lambda \end{pmatrix}$ und $B = \begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix}$ nicht ähnlich sind. *Hinweis: Sind A und B zwei ähnliche Matrizen, dann sind auch $A - \lambda I$ und $B - \lambda I$ ähnliche Matrizen.*

Übungen zu “Lineare Algebra und Geometrie 1”

Sommersemester 2012

Stefan Haller

1. Sei \mathbb{K} ein Körper, $n \in \mathbb{N}$ und bezeichne $V := F(M_{n \times n}(\mathbb{K}), \mathbb{K})$ den Vektorraum aller \mathbb{K} -wertigen Funktionen auf $M_{n \times n}(\mathbb{K})$. Zeige, dass folgende Teilmengen Teilräume von V bilden:

- (a) $\{\delta \in V : \delta \text{ ist linear in der } k\text{-ten Spalte}\}$.
- (b) $\{\delta \in V : \delta \text{ ist multilinear, d.h. linear in jeder Spalte}\}$
- (c) $\{\delta \in V : \text{stimmen die } k\text{-te und } l\text{-te Spalte von } A \text{ überein, so gilt } \delta(A) = 0\}$.
- (d) $\{\delta \in V : \text{stimmen zwei benachbarte Sp. von } A \text{ überein, so gilt } \delta(A) = 0\}$.
- (e) $\{\delta \in V : \text{stimmen zwei Spalten von } A \text{ überein, so gilt } \delta(A) = 0\}$.
- (f) $\{\delta \in V : \text{sind die Spalten von } A \text{ linear abhängig, so gilt } \delta(A) = 0\}$.
- (g) $\{\delta \in V : \delta \text{ ist multilinear und alternierend}\}$

Hinweis: Für jede Teilmenge $X \subseteq M_{n \times n}(\mathbb{K})$ ist $\{\delta \in V \mid \forall x \in X : \delta(x) = 0\}$ ein Teilraum von V .

2. Sei \mathbb{K} ein Körper in dem $2 \neq 0$ gilt. Weiters sei $n \in \mathbb{N}$ und $\delta: M_{n \times n}(\mathbb{K}) \rightarrow \mathbb{K}$ multilinear. Zeige, dass folgende Aussagen äquivalent sind:

- (a) Sind zwei Spalten von A gleich, so gilt $\delta(A) = 0$.
 - (b) Bei Vertauschung zweier Spalten von A wechselt $\delta(A)$ das Vorzeichen.
- Gib eine multilineare Abbildung $\delta: M_{2 \times 2}(\mathbb{Z}_2) \rightarrow \mathbb{Z}_2$ an, die (b) aber nicht (a) erfüllt.

3. Berechne folgende Determinanten:

$$\begin{vmatrix} 3 & 4 \\ 5 & 6 \end{vmatrix}, \quad \begin{vmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \end{vmatrix}, \quad \begin{vmatrix} 0 & 1 & 6 & 11 \\ 1 & 0 & 4 & 13 \\ 2 & -1 & 2 & 17 \\ 3 & -1 & 9 & 0 \end{vmatrix}$$

4. Für $\alpha \in \mathbb{R}$ berechne die Determinante:

$$\begin{vmatrix} \cos \alpha & -\sin \alpha & & & \\ \sin \alpha & \cos \alpha & & & \\ & & 3 & & \\ & & & 5 & \\ & & & & 7 \end{vmatrix}$$

5. Berechne die Determinante folgender $(n \times n)$ -Matrix:

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

6. Zeige $\det(\lambda A) = \lambda^n \det(A)$, für alle $A \in M_{n \times n}(\mathbb{K})$ und $\lambda \in \mathbb{K}$. Gib zwei Matrizen A und B an, für die $\det(A + B) \neq \det(A) + \det(B)$ gilt.

7. Seien $n, m \in \mathbb{N}$, $A \in M_{n \times n}(\mathbb{K})$, $B \in M_{n \times m}(\mathbb{K})$, $C \in M_{m \times n}(\mathbb{K})$ und $D \in M_{m \times m}(\mathbb{K})$. Zeige:

$$\det \begin{pmatrix} A & B \\ 0 & D \end{pmatrix} = \det(A) \det(D) = \det \begin{pmatrix} A & 0 \\ C & D \end{pmatrix}.$$

8. Zeige

$$\det \begin{pmatrix} * & * & * & * & * \\ * & * & * & * & * \\ 0 & 0 & 0 & * & * \\ 0 & 0 & 0 & * & * \\ 0 & 0 & 0 & * & * \end{pmatrix} = 0,$$

wobei die Sternchen beliebige Eintragungen bezeichnen.

9. Seien $n, m \in \mathbb{N}$, $A \in M_{n \times n}(\mathbb{K})$, $B \in M_{n \times m}(\mathbb{K})$, $C \in M_{m \times n}(\mathbb{K})$ und $D \in M_{m \times m}(\mathbb{K})$. Berechne

$$\det \begin{pmatrix} A & B \\ C & 0 \end{pmatrix} \quad \text{und} \quad \det \begin{pmatrix} 0 & B \\ C & D \end{pmatrix}.$$

Hinweis: Dies lässt sich auf Aufgabe 7 zurückführen.

10. Seien $a, b, c \in \mathbb{K}$ und betrachte folgende $(n \times n)$ -Determinante:

$$D_n := \begin{vmatrix} a & b & & & \\ c & a & b & & \\ & c & a & \ddots & \\ & & \ddots & \ddots & b \\ & & & c & a \end{vmatrix}$$

Zeige, dass die Rekursionsgleichung

$$D_n = aD_{n-1} - bcD_{n-2}$$

gilt und berechne damit die $(n \times n)$ -Determinante:

$$\begin{vmatrix} 2 & -1 & & & \\ -1 & 2 & -1 & & \\ & -1 & 2 & \ddots & \\ & & \ddots & \ddots & -1 \\ & & & -1 & 2 \end{vmatrix}$$

11. Löse das Gleichungssystem

$$\begin{pmatrix} 3 & 2 & 1 \\ 1 & -1 & 2 \\ 4 & 2 & 3 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$$

einmal mit der Cramer'schen Regel und einmal mit Zeilenumformungen.

12. Löse das Gleichungssystem

$$\begin{pmatrix} 1 & 2 & 3 & 1 \\ 0 & -1 & -2 & -4 \\ -1 & -2 & -2 & 4 \\ 0 & 1 & 2 & 3 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 17 \\ -11 \\ -10 \\ 10 \end{pmatrix}$$

einmal mit der Cramer'schen Regel und einmal mit Zeilenumformungen.

13. Berechne die Inverse der Matrix

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

einmal mit Hilfe der Formel in Bemerkung V.2.5 und einmal mit Zeilenumformungen.

14. Mit der Notation von Satz V.2.3 zeige, dass auch $\tilde{A}A = \det(A)I_n$ gilt, selbst wenn A nicht invertierbar ist.

15. Sei A eine $(n \times n)$ -Matrix mit ganzzahligen Eintragungen und $\det(A) = \pm 1$. Zeige, dass dann auch die Inverse von A nur ganzzahlige Eintragungen besitzt. Zeige auch umgekehrt: haben A und die Inverse von A nur ganzzahlige Eintragungen, dann muss schon $\det(A) = \pm 1$ gelten. Gib auch eine invertierbare Matrix mit ganzzahligen Eintragungen an, deren Inverse nicht ganzzahlig ist.

16. Sei $A \in M_{n \times n}(\mathbb{K})$ eine invertierbare Matrix. Verwende die Cramer'sche Regel um die Gleichungssysteme $Ax = e_i$ zu lösen, $1 \leq i \leq n$. Leite daraus erneut die Formel $A^{-1} = \det(A)^{-1}\tilde{A}$ aus Satz V.2.3 für invertierbare A her.

17. Zeige, dass die Permutationsgruppe \mathfrak{S}_n genau $n!$ Elemente besitzt. Wieviele Elemente hat die Untergruppe $A_n = \{\sigma \in \mathfrak{S}_n \mid \text{sgn}(\sigma) = 1\}$?

18. Schreibe alle Elemente von \mathfrak{S}_4 an, bestimme ihre Vorzeichen (Signum) und schreibe damit die Leibniz'sche Formel für (4×4) -Matrizen an.

19. Berechne die (4×4) -Determinante

$$\begin{vmatrix} 0 & 0 & 1 & 2 \\ 0 & 0 & 3 & 4 \\ 1 & 2 & 3 & 0 \\ 2 & 3 & 0 & 8 \end{vmatrix},$$

einmal mit Leibniz' Formel und einmal mit Zeilen- bzw. Spaltenumformungen.

20. Zeige, dass

$$\phi: \mathfrak{S}_n \rightarrow \mathrm{GL}_n(\mathbb{K}), \quad \phi(\sigma) := (e_{\sigma(1)} | e_{\sigma(2)} | \cdots | e_{\sigma(n)}),$$

einen Gruppenhomomorphismus bildet, d.h. es gilt $\phi(\sigma \circ \sigma') = \phi(\sigma)\phi(\sigma')$, für je zwei Permutationen $\sigma, \sigma' \in \mathfrak{S}_n$. SchlieÙe daraus

$$\mathrm{sgn}(\sigma) = \det(e_{\sigma(1)} | e_{\sigma(2)} | \cdots | e_{\sigma(n)}),$$

für jede Permutation $\sigma \in \mathfrak{S}_n$. *Hinweis: Die rechte Seite definiert einen Homomorphismus $\mathfrak{S}_n \rightarrow \{1, -1\}$, der alle Transpositionen auf -1 abbildet, vgl. Lemma V.3.2.* Dies ermöglicht folgende direkte Herleitung der Leibniz'schen Formel, erkläre dabei jedes Gleichheitszeichen:

$$\begin{aligned} \det(A) &= \sum_{i_1=1}^n \cdots \sum_{i_n=1}^n A_{1,i_1} \cdots A_{n,i_n} \det(e_{i_1} | \cdots | e_{i_n}) \\ &= \sum_{\sigma \in \mathfrak{S}_n} A_{1,\sigma(1)} \cdots A_{n,\sigma(n)} \det(e_{\sigma(1)} | \cdots | e_{\sigma(n)}) = \sum_{\sigma \in \mathfrak{S}_n} \mathrm{sgn}(\sigma) A_{1,\sigma(1)} \cdots A_{n,\sigma(n)} \end{aligned}$$

21. Zeige, dass für jede Matrix $A \in M_{n \times n}(\mathbb{K})$ die Gleichung

$$\sum_{\sigma \in \mathfrak{S}_n} \mathrm{sgn}(\sigma) A_{1,\sigma(1)} \cdots A_{n,\sigma(n)} = \sum_{\sigma \in \mathfrak{S}_n} \mathrm{sgn}(\sigma) A_{\sigma(1),1} \cdots A_{\sigma(n),n}$$

gilt. Leite daraus und der Leibniz'schen Formel erneut $\det(A) = \det(A^t)$ her.

22. Zeige, dass $\mathrm{GL}_n(\mathbb{R})$ offen in $M_{n \times n}(\mathbb{R}) = \mathbb{R}^{n^2}$ ist. Zeige, dass $\mathrm{SL}_n(\mathbb{R})$ abgeschlossen in $M_{n \times n}(\mathbb{R}) = \mathbb{R}^{n^2}$ ist. *Hinweis: Die Determinantenfunktion ist stetig.*

23. Seien V ein n -dimensionaler \mathbb{K} -Vektorraum, $\varphi \in \mathrm{end}(V)$, B eine geordnete Basis von V und $[\varphi]_{BB} \in M_{n \times n}(\mathbb{K})$ die Matrix von φ bezüglich B . Zeige, dass $\mathrm{tr}([\varphi]_{BB})$ nicht von der Basis B abhängt und daher

$$\mathrm{tr}: \mathrm{end}(V) \rightarrow \mathbb{K}, \quad \mathrm{tr}(\varphi) = \mathrm{tr}([\varphi]_{BB}).$$

wohldefiniert ist. Zeige auch:

- (a) Die Spur ist linear, d.h. $\mathrm{tr}(\varphi_1 + \varphi_2) = \mathrm{tr}(\varphi_1) + \mathrm{tr}(\varphi_2)$ und $\mathrm{tr}(\lambda\varphi) = \lambda \mathrm{tr}(\varphi)$.
- (b) $\mathrm{tr}(\psi \circ \varphi) = \mathrm{tr}(\varphi \circ \psi)$.
- (c) $\mathrm{tr}(\varphi^t) = \mathrm{tr}(\varphi)$.
- (d) $\mathrm{tr}(\mathrm{id}_V) = \dim(V)$.

für $\varphi, \psi, \varphi_1, \varphi_2 \in \mathrm{end}(V)$ und $\lambda \in \mathbb{K}$. *Hinweis: Gehe wie in Korollar V.4.2 vor.*

24. Seien W und W' zwei komplementäre Teilräume eines endlich-dimensionalen Vektorraums V , d.h. $V = W \oplus W'$. Weiters bezeichne $\pi: V \rightarrow V$ die Projektion auf W längs W' und $\sigma: V \rightarrow V$ die Spiegelung an W längs W' . Berechne $\mathrm{tr}(\pi)$ und $\det(\sigma)$. Berechne auch $\det(\lambda \mathrm{id}_V)$, wobei $\lambda \in \mathbb{K}$.

25. Sei V ein endlich-dimensionaler \mathbb{K} -Vektorraum, W ein Teilraum von V und $\varphi: V \rightarrow V$ linear, sodass $\varphi(W) \subseteq W$. Zeige

$$\det(\varphi) = \det(\varphi|_W) \det(\bar{\varphi}),$$

wobei $\varphi|_W: W \rightarrow W$ die Einschränkung von φ bezeichnet, und $\bar{\varphi}$ die auf dem Quotientenraum V/W induzierte lineare Abbildung $\bar{\varphi}: V/W \rightarrow V/W$, $\bar{\varphi}([v]) = [\varphi(v)]$ bezeichnet. *Hinweis: Dies lässt sich durch geschickte Wahl einer Basis von V auf Aufgabe 7 zurückführen.*

26. Sei V ein endlich-dimensionaler reeller Vektorraum. Zeige, dass

$$\mathrm{GL}^+(V) := \{\varphi \in \mathrm{end}(V) : \det(\varphi) > 0\}$$

eine Untergruppe von $\mathrm{GL}(V)$ bildet.

27. Sei V ein endlich-dimensionaler Vektorraum über \mathbb{K} . Zeige, dass der Homomorphismus $\det: \mathrm{GL}(V) \rightarrow \mathbb{K} \setminus \{0\}$ surjektiv ist.

28. Betrachte die beiden geordneten Basen $B = (b_1, b_2, b_3)$ und $B' = (b'_1, b'_2, b'_3)$ von \mathbb{R}^3 , wobei:

$$b_1 = \begin{pmatrix} 2 \\ 0 \\ 0 \end{pmatrix}, \quad b_2 = \begin{pmatrix} 3 \\ 4 \\ 0 \end{pmatrix}, \quad b_3 = \begin{pmatrix} 5 \\ 6 \\ 7 \end{pmatrix}, \quad b'_1 = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}, \quad b'_2 = \begin{pmatrix} 0 \\ 0 \\ 6 \end{pmatrix}, \quad b'_3 = \begin{pmatrix} 0 \\ 7 \\ 8 \end{pmatrix}$$

Sind B und B' gleichorientiert? Welche dieser Basen ist bezüglich der Standardorientierung auf \mathbb{R}^3 positiv orientiert?

29. Zeige, dass für jedes $A \in M_{n \times n}(\mathbb{K})$ folgende Zahlen übereinstimmen:

- (a) $\mathrm{rank}(A)$
- (b) das größte k , für das es eine invertierbare $(k \times k)$ -Untermatrix von A gibt
- (c) das größte k , für das es eine $(k \times k)$ -Untermatrix von A gibt, deren Determinante nicht verschwindet

Dabei verstehen wir unter einer $(k \times k)$ -Untermatrix von A jede Matrix der Form

$$\begin{pmatrix} a_{i_1 j_1} & a_{i_1 j_2} & \cdots & a_{i_1 j_k} \\ a_{i_2 j_1} & a_{i_2 j_2} & \cdots & a_{i_2 j_k} \\ \vdots & \vdots & \ddots & \vdots \\ a_{i_k j_1} & a_{i_k j_2} & \cdots & a_{i_k j_k} \end{pmatrix}$$

wobei $1 \leq i_1 < i_2 < \cdots < i_k \leq n$ und $1 \leq j_1 < j_2 < \cdots < j_k \leq n$ und a_{ij} die Eintragung von A in der i -ten Zeile und j -ten Spalte bezeichnet.

30. Sei V ein endlich-dimensionaler komplexer Vektorraum und $\varphi: V \rightarrow V$ (komplex) linear. Es bezeichne $V_{\mathbb{R}}$ den zugrundeliegenden reellen Vektorraum und $\varphi_{\mathbb{R}}: V_{\mathbb{R}} \rightarrow V_{\mathbb{R}}$ dieselbe Abbildung, aufgefasst als reell lineare Abbildung. Zeige

$$\det(\varphi_{\mathbb{R}}) = |\det(\varphi)|^2,$$

und schlieÙe daraus, dass im invertierbaren Fall $\varphi_{\mathbb{R}}$ orientierungsbewahrend ist. *Hinweis:* Ist $B = (b_1, \dots, b_n)$ eine Basis des komplexen Vektorraums V , dann bildet $\tilde{B} = (b_1, \mathbf{i}b_1, b_2, \mathbf{i}b_2, \dots, b_n, \mathbf{i}b_n)$ eine Basis des reellen Vektorraums $V_{\mathbb{R}}$, und es gilt

$$[\varphi_{\mathbb{R}}]_{\tilde{B}\tilde{B}} = \left(\begin{array}{cc|ccc} x_{11} & -y_{11} & \cdots & x_{1n} & -y_{1n} \\ y_{11} & x_{11} & \cdots & y_{1n} & x_{1n} \\ \vdots & \vdots & & \vdots & \vdots \\ \hline x_{n1} & -y_{n1} & \cdots & x_{nn} & -y_{nn} \\ y_{n1} & x_{n1} & \cdots & y_{nn} & x_{nn} \end{array} \right), \text{ wobei } [\varphi]_{BB} = \begin{pmatrix} z_{11} & \cdots & z_{1n} \\ \vdots & & \vdots \\ z_{n1} & \cdots & z_{nn} \end{pmatrix}$$

und $z_{ij} = x_{ij} + \mathbf{i}y_{ij}$. Die gewünschte Gleichung lässt sich mittels Induktion nach n zeigen, für den Induktionsschritt betrachte zunächst den Fall $z_{21} = z_{31} = \dots = z_{n1} = 0$. Welcher Zusammenhang besteht zwischen $\text{tr}(\varphi)$ und $\text{tr}(\varphi_{\mathbb{R}})$?

31. Für jede Permutation $\sigma \in \mathfrak{S}_n$ bezeichne $S_{\sigma} = (e_{\sigma(1)} | \dots | e_{\sigma(n)})$ die entsprechende Permutationsmatrix.

a) Zeige:

$$(S_{\sigma})^{-1} \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix} S_{\sigma} = \begin{pmatrix} \lambda_{\sigma(1)} & & \\ & \ddots & \\ & & \lambda_{\sigma(n)} \end{pmatrix}.$$

b) Sei D eine $(n \times n)$ -Diagonalmatrix mit paarweise verschiedenen Diagonaleinträgen und A eine weitere $(n \times n)$ -Matrix. Zeige, dass $AD = DA$ genau dann gilt, wenn A eine Diagonalmatrix ist. Zeige anhand eines Beispiels, dass die Voraussetzung an die Diagonaleinträge von D wirklich notwendig ist.

c) Sei D eine $(n \times n)$ -Diagonalmatrix mit paarweise verschiedenen Diagonaleinträgen und S eine invertierbare Matrix, sodass auch $S^{-1}DS$ eine Diagonalmatrix bildet. Zeige, dass dann S von der Form $S = AS_{\sigma}$ sein muss, wobei A eine invertierbare $(n \times n)$ -Diagonalmatrix bezeichnet und $\sigma \in \mathfrak{S}_n$. *Hinweis:* Zeige zunächst mit Hilfe von a), dass o.B.d.A. $S^{-1}DS = D$ angenommen werden kann, und verwende dann b).

32. Bestimme alle Eigenwerte und Eigenräume folgender reeller Matrizen:

$$A = \begin{pmatrix} 2 & 1 \\ 4 & 5 \end{pmatrix}, \quad B = \begin{pmatrix} 5 & -1 \\ 4 & 1 \end{pmatrix}.$$

Welche dieser Matrizen sind diagonalisierbar? Gib im diagonalisierbaren Fall Matrizen S an, sodass $S^{-1}AS$ bzw. $S^{-1}BS$ Diagonalmatrizen bilden.

33. Wie im vorangehenden Beispiel nun für die reellen Matrizen:

$$A = \begin{pmatrix} 1 & 0 & 0 \\ 3 & 20 & 30 \\ -2 & -12 & -18 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & -3 & 2 \\ 0 & -8 & 6 \\ 0 & -15 & 11 \end{pmatrix}, \quad C = \begin{pmatrix} 2 & 0 & 1 \\ 0 & 2 & 0 \\ 0 & 1 & 2 \end{pmatrix}$$

34. Bestimme alle Eigenwerte und Eigenräume der Matrix

$$A = \begin{pmatrix} \lambda & 1 & & \\ & \lambda & \ddots & \\ & & \ddots & 1 \\ & & & \lambda \end{pmatrix}$$

wobei $\lambda \in \mathbb{K}$. Ist A diagonalisierbar?

35. Sei $q \in \mathbb{K}[z]$ ein Polynom und D eine Dreiecksmatrix,

$$D = \begin{pmatrix} \lambda_1 & * & \cdots & * \\ & \lambda_2 & \ddots & \vdots \\ & & \ddots & * \\ & & & \lambda_n \end{pmatrix}. \quad \text{Zeige} \quad q(D) = \begin{pmatrix} q(\lambda_1) & * & \cdots & * \\ & q(\lambda_2) & \ddots & \vdots \\ & & \ddots & * \\ & & & q(\lambda_n) \end{pmatrix}.$$

36. Sei V ein endlich dimensionaler \mathbb{K} -Vektorraum, $\varphi: V \rightarrow V$ eine diagonalisierbare lineare Abbildung und $q \in \mathbb{K}[z]$ ein Polynom. Zeige, dass dann auch $q(\varphi)$ diagonalisierbar ist und bestimme das Spektrum von $q(\varphi)$.

37. Erkläre warum die Matrix

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

über dem Körper $\mathbb{K} = \mathbb{Q}$ nicht diagonalisierbar ist. Zeige, dass sie über $\mathbb{K} = \mathbb{R}$ sehr wohl diagonalisierbar ist, bestimme ihre Eigenwert und Eigenräume sowie eine invertierbare Matrix S , sodass $S^{-1}AS$ Diagonalgestalt hat.

38. Sei V ein endlich dimensionaler \mathbb{K} -Vektorraum, $\varphi: V \rightarrow V$ linear, $\lambda \in \mathbb{K}$ ein Eigenwert von φ und $E_\lambda = \ker(\varphi - \lambda \text{id}_V)$ der entsprechende Eigenraum. Darüber hinaus, sei $\psi: V \rightarrow V$ eine weitere lineare Abbildung, die mit φ kommutiert, d.h. $\varphi\psi = \psi\varphi$. Zeige, dass ψ die Eigenräume von φ invariant lässt, d.h. $\psi(E_\lambda) \subseteq E_\lambda$.

39. Betrachte die Folge $0, 1, 2, 7, 20, 61, 182, \dots$. Genauer sei

$$x_0 = 0, \quad x_1 = 1, \quad x_n = 2x_{n-1} + 3x_{n-2}, \quad n \geq 2.$$

Bestimme eine explizite Formel für x_n . Gehe dabei wie in Beispiel VI.1.26 vor.

40. Sei $A \in M_{n \times n}(\mathbb{K})$.

- Welcher Zusammenhang besteht zwischen den Eigenwerten und Eigenräumen von A und denen von μA , wobei $0 \neq \mu \in \mathbb{K}$?
- Welcher Zusammenhang besteht zwischen den Eigenwerten und Eigenräumen von A und denen von A^{-1} , falls A invertierbar ist.
- Welcher Zusammenhang besteht zwischen den Eigenwerten von A und denen von A^t ?

Gehe jeweils auch auf die geometrischen und algebraischen Vielfachheiten ein.

41. Sei $A \in M_{n \times n}(\mathbb{C})$. Zeige

$$\left. \frac{d}{dt} \right|_0 \det(I_n + tA) = \operatorname{tr}(A).$$

Hinweis: Es existiert $S \in \operatorname{GL}_n(\mathbb{C})$, sodass $S^{-1}AS$ eine obere Dreiecksmatrix ist. Erkläre warum obige Formel auch für reelle Matrizen $A \in M_{n \times n}(\mathbb{R})$ richtig bleibt.

42. Zeige, dass jedes Polynom der Form $p = p_0 + p_1z + \dots + p_{n-1}z^{n-1} + (-1)^n z^n$, wobei $p_i \in \mathbb{K}$, als charakteristisches Polynom einer Matrix $A \in M_{n \times n}(\mathbb{K})$ auftritt. Berechne dazu das charakteristische Polynom der Matrix

$$A = \begin{pmatrix} 0 & \cdots & 0 & (-1)^{n-1}p_0 \\ 1 & \ddots & \vdots & \vdots \\ & \ddots & 0 & (-1)^{n-1}p_{n-2} \\ & & 1 & (-1)^{n-1}p_{n-1} \end{pmatrix}$$

mittels Induktion nach n .

43. Zeige, dass $\operatorname{GL}_n(\mathbb{R})$ dicht in $M_{n \times n}(\mathbb{R}) = \mathbb{R}^{n^2}$ liegt, d.h. in jeder Umgebung einer Matrix $A \in M_{n \times n}(\mathbb{R})$ gibt es invertierbare Matrizen. Betrachte dazu die Familie von Matrizen $A - t$, $t \in \mathbb{R}$, und zeige, dass es beliebig kleine t gibt für die $A - t$ invertierbar ist.

44. Sei p eine Primzahl. Zeige, dass der Körper \mathbb{Z}_p nicht algebraisch abgeschlossen ist. *Hinweis: Betrachte das Polynom $q = 1 + (z - 1)(z - 2) \cdots (z - p)$.*

45. Sei V ein endlich dimensionaler \mathbb{K} -Vektorraum und $\pi: V \rightarrow V$ ein Projektor, d.h. $\pi^2 = \pi$. Bestimme alle Eigenwerte und Eigenräume sowie das charakteristische Polynom von π . Gib jeweils auch die geometrische und algebraische Vielfachheit der Eigenwerte an.

46. Sei $I \subseteq \mathbb{R}$ ein offenes Intervall, betrachte den Vektorraum der glatten Funktionen, $V := C^\infty(I, \mathbb{R})$, und bezeichne $D: V \rightarrow V$, $D(f) := f'$, den Ableitungsoperator. Für $\lambda \in \mathbb{R}$ sei $f_\lambda \in V$, $f_\lambda(x) := e^{\lambda x}$. Zeige, dass f_λ Eigenvektor von D ist und bestimme den Eigenwert.

47. Sei $P \in M_{n \times n}(\mathbb{K}[z])$ eine Matrix von Polynomen. Zeige, dass $\det(P)$ linear in jeder Spalte ist, d.h. für beliebige Polynome $p_{ij}, \tilde{p}_{ij}, q \in \mathbb{K}[z]$ gilt

$$\det \begin{pmatrix} p_{11} & \cdots & p_{1j} + \tilde{p}_{1j} & \cdots & p_{1n} \\ \vdots & & \vdots & & \vdots \\ p_{n1} & \cdots & p_{nj} + \tilde{p}_{nj} & \cdots & p_{nn} \end{pmatrix} = \det \begin{pmatrix} p_{11} & \cdots & p_{1j} & \cdots & p_{1n} \\ \vdots & & \vdots & & \vdots \\ p_{n1} & \cdots & p_{nj} & \cdots & p_{nn} \end{pmatrix} + \det \begin{pmatrix} p_{11} & \cdots & \tilde{p}_{1j} & \cdots & p_{1n} \\ \vdots & & \vdots & & \vdots \\ p_{n1} & \cdots & \tilde{p}_{nj} & \cdots & p_{nn} \end{pmatrix}$$

und

$$\det \begin{pmatrix} p_{11} & \cdots & qp_{1j} & \cdots & p_{1n} \\ \vdots & & \vdots & & \vdots \\ p_{n1} & \cdots & qp_{nj} & \cdots & p_{nn} \end{pmatrix} = q \det \begin{pmatrix} p_{11} & \cdots & p_{1j} & \cdots & p_{1n} \\ \vdots & & \vdots & & \vdots \\ p_{n1} & \cdots & p_{nj} & \cdots & p_{nn} \end{pmatrix}$$

Zeige weiters $\det(P^t) = \det(P)$, vgl. Aufgabe 21, und schließe daraus, dass $\det(P)$ auch linear in jeder Spalte ist.

48. Sei $P = (p_1 | \cdots | p_n) \in M_{n \times n}(\mathbb{K}[z])$ eine Matrix von Polynomen. Zeige $\det(p_{\sigma(1)} | \cdots | p_{\sigma(n)}) = \text{sign}(\sigma) \det(P) \in \mathbb{K}[z]$, für jede Permutation $\sigma \in \mathfrak{S}_n$.

49. Seien $P, Q \in M_{n \times n}(\mathbb{K}[z])$ zwei Matrizen, deren Eintragungen Polynome sind. Zeige $\det(PQ) = \det(P) \det(Q) \in \mathbb{K}[z]$ auf zwei Arten: Für unendliche Körper unter Verwendung von $(\det(P))(\lambda) = \det(P(\lambda))$, $\lambda \in \mathbb{K}$, und dann für beliebige Körper. *Anleitung für den zweiten Teil:*

- 1.) Zeige, dass $\det(PQ)$ und $\det(P) \det(Q)$ beide linear in jeder Spalte von Q sind, siehe Aufgabe 47, und schließe daraus, dass es o.B.d.A. genügt Matrizen der Form $Q = (e_{i_1} | \cdots | e_{i_n})$ zu betrachten.
- 2.) Zeige, dass $\det(PQ) = 0 = \det(Q)$, falls Q zwei gleiche Spalten besitzt, und schließe daraus, dass es o.B.d.A. genügt Permutationsmatrizen $Q = (e_{\sigma(1)} | \cdots | e_{\sigma(n)})$ zu betrachten, wobei $\sigma \in \mathfrak{S}_n$.
- 3.) Verwende Aufgabe 48 um den Beweis abzuschließen.

50. Zeige, dass für jede Matrix $A \in M_{2 \times 2}(\mathbb{K})$ die Relation

$$\det(A) = \frac{1}{2} \left(\text{tr}(A)^2 - \text{tr}(A^2) \right)$$

gilt und analog für $A \in M_{3 \times 3}(\mathbb{K})$

$$\det(A) = \frac{1}{6} \left(\text{tr}(A)^3 - 3 \text{tr}(A) \text{tr}(A^2) + 2 \text{tr}(A^3) \right).$$

Hinweis: Es genügt dies für algebraisch abgeschlossene Körper, d.h. triangulierbare Matrizen, zu zeigen.

51. Betrachte die Permutation $\sigma \in \mathfrak{S}_n$,

$$\sigma(1) = 2, \quad \sigma(2) = 3, \quad \sigma(3) = 4, \quad \dots \quad \sigma(n-1) = n, \quad \sigma(n) = 1,$$

und die assoziierte Permutationsmatrix

$$A = (e_{\sigma(1)} | \cdots | e_{\sigma(n)}) = \begin{pmatrix} 0 & 0 & \cdots & 0 & 1 \\ 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \ddots & \vdots & \vdots \\ \vdots & \ddots & \ddots & 0 & 0 \\ 0 & \cdots & 0 & 1 & 0 \end{pmatrix} \in M_{n \times n}(\mathbb{C}),$$

Bestimme alle komplexen Eigenwerte von A sowie deren algebraische und geometrische Vielfachheit. Ist A diagonalisierbar? *Freiwillige Zusatzaufgabe: Bestimme eine Eigenbasis von A .*

52. Seien $A, B \in M_{n \times n}(\mathbb{K})$ zwei Matrizen von denen mindestens eine invertierbar ist. Zeige, dass AB und BA die selben Eigenwerte mit den gleichen algebraischen und geometrischen Vielfachheiten haben.

53. Welche der folgenden reellen Matrizen sind triangulierbar?

$$A = \begin{pmatrix} 2 & 5 \\ -1 & -2 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & 0 \\ 2 & 3 \end{pmatrix}, \quad C = \begin{pmatrix} 1 & -2 \\ 1 & 4 \end{pmatrix}, \quad D = \begin{pmatrix} 5 & 1 \\ -9 & -1 \end{pmatrix}.$$

Gib im triangulierbaren Fall jeweils invertierbare Matrizen S an, sodass $S^{-1}AS$, $S^{-1}BS$, etc. obere Dreiecksgestalt besitzt.

54. Wie in der vorangehenden Aufgabe nun mit den reellen Matrizen:

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

$$C = \begin{pmatrix} -4 & 1 & -2 \\ 3 & 2 & 1 \\ 18 & -3 & 8 \end{pmatrix}, \quad D = \begin{pmatrix} -3 & -8 & -2 \\ 2 & 5 & 1 \\ 8 & 12 & 5 \end{pmatrix}.$$

55. Sei V ein n -dimensionaler Vektorraum und $\varphi: V \rightarrow V$ linear. Zeige, dass die folgenden Aussagen äquivalent sind:

- (a) φ ist triangulierbar, d.h. es existiert eine Basis B von V , sodass $[\varphi]_{BB}$ obere Dreiecksgestalt hat.
- (b) Es existieren Teilräume $\{0\} = V_0 \subseteq V_1 \subseteq V_2 \subseteq \dots \subseteq V_n = V$ mit $\dim(V_i) = i$ und $\varphi(V_i) \subseteq V_i$, für jedes $i = 1, \dots, n$.

56. Sei V ein n -dimensionaler Vektorraum und $\varphi: V \rightarrow V$ linear. Zeige, dass φ genau dann triangulierbar ist, wenn $\varphi^t: V^* \rightarrow V^*$ triangulierbar ist. Schließe daraus, dass eine Matrix $A \in M_{n \times n}(\mathbb{K})$ genau dann triangulierbar ist, wenn die transponierte Matrix $A^t \in M_{n \times n}(\mathbb{K})$ triangulierbar ist.

57. Sei $\lambda \in \mathbb{K}$. Für jede der Matrizen

$$\begin{pmatrix} \lambda & 1 & & & \\ & \lambda & 1 & & \\ & & \lambda & 1 & \\ & & & \lambda & 1 \\ & & & & \lambda \end{pmatrix}, \quad \begin{pmatrix} \lambda & 1 & & & \\ & \lambda & 1 & & \\ & & \lambda & 0 & \\ & & & \lambda & 1 \\ & & & & \lambda \end{pmatrix}, \quad \begin{pmatrix} \lambda & 1 & & & \\ & \lambda & 0 & & \\ & & \lambda & 1 & \\ & & & \lambda & 1 \\ & & & & \lambda \end{pmatrix}$$

bestimme die Dimensionen der Teilräume

$$E_\lambda = \ker(A - \lambda) \subseteq \ker((A - \lambda)^2) \subseteq \dots \subseteq \ker((A - \lambda)^N) = \tilde{E}_\lambda$$

sowie das minimale N für das $\ker((A - \lambda)^N) = \tilde{E}_\lambda$ gilt, wobei A jeweils eine der Matrizen bezeichnet.

58. Für $\varepsilon \in \mathbb{K}$ zeige:

$$S^{-1} \begin{pmatrix} \lambda & 1 & & \\ & \lambda & \ddots & \\ & & \ddots & 1 \\ & & & \lambda \end{pmatrix} S = \begin{pmatrix} \lambda & \varepsilon & & \\ & \lambda & \ddots & \\ & & \ddots & \varepsilon \\ & & & \lambda \end{pmatrix}, \quad \text{wobei} \quad S = \begin{pmatrix} 1 & & & \\ & \varepsilon & & \\ & & \ddots & \\ & & & \varepsilon^{n-1} \end{pmatrix}$$

59. Verifiziere den Satz von Caley–Hamilton anhand folgender Matrix,

$$A := \begin{pmatrix} 5 & 0 & 4 \\ -3 & 2 & -7 \\ 0 & 0 & 3 \end{pmatrix},$$

d.h. bestimme das charakteristische Polynom p von A und berechne $p(A)$.

60. Sei $A \in M_{n \times n}(\mathbb{K})$ eine Matrix, deren charakteristisches Polynom in Linearfaktoren zerfällt, d.h. $p = (\lambda_1 - z)^{m_1} \cdots (\lambda_k - z)^{m_k}$, wobei $\lambda_1, \dots, \lambda_k$ die verschiedenen Eigenwerte und m_1, \dots, m_k ihre algebraischen Vielfachheiten bezeichnen. Beweise erneut den Satz von Caley–Hamilton, d.h. $p(A) = 0$, nun in Matrixschreibweise. *Anleitung:*

(1) Nach Satz VI.3.7 über die Primärzerlegung existiert $S \in \text{GL}_n(\mathbb{K})$, sodass

$$B := S^{-1}AS = \begin{pmatrix} B_1 & & \\ & \ddots & \\ & & B_k \end{pmatrix}, \quad B_i = \begin{pmatrix} \lambda_i & * & \cdots & * \\ & \lambda_i & \ddots & \vdots \\ & & \ddots & * \\ & & & \lambda_i \end{pmatrix} \in M_{m_i \times m_i}(\mathbb{K}).$$

(2) Wie sehen die Diagonaleinträge des i -ten Diagonalblocks von $(B - \lambda_i I_n)$ aus?

(3) Wie sieht der i -te Diagonalblock von $(B - \lambda_i I_n)^{m_i}$ aus?

(4) SchlieÙe daraus, dass $p(B) = (\lambda_1 I_n - B)^{m_1} \cdots (\lambda_k I_n - B)^{m_k} = 0$.

(5) Folgere $p(A) = 0$.

61. Sei V ein endlich dimensionaler \mathbb{K} -Vektorraum und $\varphi: V \rightarrow V$ linear. Weiters sei $\psi: V \rightarrow V$ linear und kommutiere mit φ , d.h. $\varphi\psi = \psi\varphi$. Zeige, $\psi(\tilde{E}_\lambda) \subseteq \tilde{E}_\lambda$, wobei \tilde{E}_λ den verallgemeinerten Eigenraum zu einem Eigenwert λ von φ bezeichnet.

62. Seien $\varphi, \psi: V \rightarrow V$ zwei kommutierende nilpotente lineare Abbildungen. Zeige, dass dann auch $\varphi + \psi$ nilpotent ist. Gib zwei (nicht kommutierende) nilpotente Abbildungen an, deren Summe nicht nilpotent ist.

63. Sei V ein endlich dimensionaler \mathbb{K} -Vektorraum und $\varphi: V \rightarrow V$ eine diagonalisierbare lineare Abbildung. Weiters sei $V = W \oplus W'$ eine invariante Zerlegung, d.h. $\varphi(W) \subseteq W$ und $\varphi(W') \subseteq W'$. Zeige, dass dann auch die Einschränkungen $\varphi|_W: W \rightarrow W$ und $\varphi|_{W'}: W' \rightarrow W'$ diagonalisierbar sind.

64. Zeige, dass die einzige diagonalisierbare und nilpotente lineare Abbildung die Nullabbildung ist.

65. Seien $\varphi, \psi: V \rightarrow V$ zwei kommutierende diagonalisierbare lineare Abbildungen. Zeige, dass φ und ψ gleichzeitige diagonalisierbar sind, d.h. es existiert eine Basis B von V , sodass $[\varphi]_{BB}$ und $[\psi]_{BB}$ Diagonalmatrizen sind.

66. Betrachte die Matrix

$$A = \begin{pmatrix} 3 & -1 & 0 & 3 \\ 0 & 0 & 0 & 8 \\ 0 & 1 & 2 & -4 \\ 0 & 0 & 1 & 4 \end{pmatrix}.$$

Bestimme eine diagonalisierbare Matrix D sowie eine nilpotente Matrix N , sodass $A = D + N$ und $DN = ND$.

67. Zeige folgende Formel für die k -te Potenz eines $(n \times n)$ -Jordanblocks:

$$\begin{pmatrix} \lambda & 1 & & & \\ & \lambda & \ddots & & \\ & & \ddots & \ddots & \\ & & & \ddots & 1 \\ & & & & \lambda \end{pmatrix}^k = \begin{pmatrix} \lambda^k & k\lambda^{k-1} & \binom{k}{2}\lambda^{k-2} & \dots & \binom{k}{n}\lambda^{k-n+1} \\ & \lambda^k & k\lambda^{k-1} & \ddots & \vdots \\ & & \lambda^k & \ddots & \binom{k}{2}\lambda^{k-2} \\ & & & \ddots & k\lambda^{k-1} \\ & & & & \lambda^k \end{pmatrix}$$

Hinweis: Induktion nach k .

68. Bestimme die Jordan'sche Normalform der Matrix

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

sowie eine invertierbare Matrix S , sodass $S^{-1}AS$ Normalform hat.

69. Bestimme die Jordan'sche Normalform der Matrix

$$A = \begin{pmatrix} 3 & 1 & 0 & 0 & 0 & 0 & 0 \\ & 3 & 1 & -2 & 0 & 0 & 4 \\ & & 3 & 0 & 1 & 0 & 0 \\ & & & 3 & 0 & 1 & 2 \\ & & & & 3 & 0 & 0 \\ & & & & & 3 & 1 \\ & & & & & & 3 \end{pmatrix}$$

sowie eine invertierbare Matrix S , sodass $S^{-1}AS$ Normalform hat.

79. Sei V ein Vektorraum über \mathbb{K} . Zeige, dass die Menge der Bilinearformen auf V bezüglich der Operationen

$$(\beta_1 + \beta_2)(v, w) := \beta_1(v, w) + \beta_2(v, w), \quad (\lambda\beta)(v, w) = \lambda\beta(v, w).$$

einen \mathbb{K} -Vektorraum bildet. Zeige weiters, dass die Menge der symmetrischen Bilinearformen ein Teilraum ist. Formuliere und beweise die analoge Aussage für Sesquilinearformen (Hermitesche Formen).

80. Sei V ein \mathbb{K} -Vektorraum und $\beta: V \times V \rightarrow \mathbb{K}$ eine Abbildung, die linear in der ersten Eintragung und symmetrisch ist, d.h. es gelte

$$\beta(v + v', w) = \beta(v, w) + \beta(v', w), \quad \beta(\lambda v, w) = \lambda\beta(v, w), \quad \beta(w, v) = \beta(v, w),$$

für alle $v, v', w \in V$ und $\lambda \in \mathbb{K}$. Zeige, dass β eine symmetrische Bilinearform ist, d.h. zeige, dass β dann auch linear in der zweiten Eintragung sein muss. Formuliere und beweise die analoge Aussage für Hermitesche Formen.

81. Bestimme die Matrix (bezüglich der Standardbasis) der symmetrischen Bilinearform auf \mathbb{R}^3 mit quadratischer Form:

$$q\left(\begin{pmatrix} x \\ y \\ z \end{pmatrix}\right) = x^2 + 2y^2 + 3z^2 + 2xy + 7xz - 6yz.$$

82. Sei $\beta: V \times V \rightarrow \mathbb{K}$ eine symmetrische Bilinearform mit assoziierter quadratischer Form $q(v) = \beta(v, v)$. Zeige, dass q der sogenannte *Parallelogrammgleichung* genügt, d.h. für alle $v, w \in V$ gilt

$$q(v + w) + q(v - w) = 2q(v) + 2q(w).$$

Formuliere und beweise die analoge Aussage für Hermitesche Formen.

83. Sei V ein \mathbb{K} -Vektorraum und $\beta: V \times V \rightarrow \mathbb{K}$ eine symmetrische Bilinearform. Zeige, dass

$$O(V, \beta) := \{\varphi \in \text{GL}(V) \mid \forall v, w \in V : \beta(\varphi(v), \varphi(w)) = \beta(v, w)\}$$

bezüglich Komposition von Abbildungen eine Gruppe bildet. Zeige weiters, dass im Fall $\text{char}(\mathbb{K}) \neq 2$,

$$O(V, \beta) = \{\varphi \in \text{GL}(V) \mid \forall v \in V : q(\varphi(v)) = q(v)\},$$

wobei $q(v) = \beta(v, v)$ die assoziierte quadratische Form bezeichnet. Formuliere und beweise die analoge Aussage für Hermitesche Formen.

84. Sei $\varphi: W \rightarrow V$ eine lineare Abbildung und $\beta: V \times V \rightarrow \mathbb{K}$ eine symmetrische Bilinearform. Zeige, dass $\varphi^*\beta$,

$$\varphi^*\beta: W \times W \rightarrow \mathbb{K}, \quad (\varphi^*\beta)(w_1, w_2) := \beta(\varphi(w_1), \varphi(w_2)),$$

eine symmetrische Bilinearform auf W definiert. Zeige:

$$(\varphi \circ \psi)^*\beta = \psi^*(\varphi^*\beta) \quad \text{und} \quad (\text{id}_V)^*\beta = \beta$$

für jede weitere lineare Abbildung $\psi: U \rightarrow W$, sowie

$$\varphi^*(\beta_1 + \beta_2) = \varphi^*\beta_1 + \varphi^*\beta_2 \quad \text{und} \quad \varphi^*(\lambda\beta) = \lambda\varphi^*\beta$$

für symmetrische Bilinearformen β, β_1, β_2 auf V und $\lambda \in \mathbb{K}$. Zeige weiters:

- (a) Ist β positiv/negativ semidefinit, dann ist $\varphi^*\beta$ positiv/negativ semidefinit.
- (b) Ist β positiv/negativ def. und φ injektiv, dann ist $\varphi^*\beta$ positiv/negativ def.
- (c) Ist β nicht-degeneriert und φ injektiv, dann ist auch $\varphi^*\beta$ nicht degeneriert.
- (d) $\iota^*\beta = \beta|_W$, wobei $\iota: W \rightarrow V$ die Inklusion eines Teilraums bezeichnet.
- (e) $O(V, \beta) = \{\varphi \in \text{GL}(V) \mid \varphi^*\beta = \beta\}$.

Formuliere und beweise analoge Eigenschaften Hermitescher Formen.

85 (Frobeniushomomorphismus). Sei \mathbb{K} ein Körper mit endlicher Charakteristik, $p = \text{char}(\mathbb{K})$. Zeige, dass die Abbildung

$$F: \mathbb{K} \rightarrow \mathbb{K}, \quad F(x) = x^p,$$

ein Körperhomomorphismus ist, d.h. es gilt

$$F(x + y) = F(x) + F(y) \quad \text{und} \quad F(xy) = F(x)F(y)$$

für alle $x, y \in \mathbb{K}$. Für die Fixpunktmenge von F zeige weiters,

$$\text{Fix}(F) := \{x \in \mathbb{K} \mid F(x) = x\} = \{0, 1, 2, \dots, p-1\}.$$

Hinweis: Nach dem kleinen Satz von Fermat gilt $n^p \equiv n \pmod{p}$, für jede Primzahl p und alle $n \in \mathbb{Z}$. Auch ist $\binom{p}{i}$ durch p teilbar, für alle $0 < i < p$.

86. Sei $\beta: V \times V \rightarrow \mathbb{K}$ eine symmetrische Bilinearform und $W := \ker(\beta)$. Zeige, dass

$$\bar{\beta}: V/W \times V/W \rightarrow \mathbb{K}, \quad \bar{\beta}([v], [v']) := \beta(v, v'),$$

eine wohldefinierte symmetrische Bilinearform definiert, die nicht-degeneriert ist.

87. Zwei symmetrische Matrizen $A, B \in M_{n \times n}^{\text{sym}}(\mathbb{K})$ werden *kongruent* genannt, falls $S \in \text{GL}_n(\mathbb{K})$ existiert, sodass $B = S^t A S$. Zeige, dass dies eine Äquivalenzrelation auf $M_{n \times n}^{\text{sym}}(\mathbb{K})$ ist. Formuliere und beweise die analoge Aussage für Hermitesche Formen.

88. Sei V ein endlich dimensionaler komplexer Vektorraum. Weiters seien β und β' zwei symmetrische Bilinearformen auf V mit $\text{rank}(\beta) = \text{rank}(\beta')$. Zeige, dass ein linearer Isomorphismus $\varphi: V \rightarrow V$ existiert, sodass $\beta' = \varphi^*\beta$, d.h. $\beta'(v, w) = \beta(\varphi(v), \varphi(w))$, für alle $v, w \in V$. *Hinweis: Satz VII.1.25.*

89. Für jede der folgenden symmetrischen Matrizen

$$A_1 = \begin{pmatrix} -1 & 0 & 0 \\ 0 & -4 & 0 \\ 0 & 0 & -9 \end{pmatrix}, \quad A_2 = \begin{pmatrix} 4 & 0 & 0 \\ 0 & \mathbf{i} & 2 \\ 0 & 2 & -4 - 7\mathbf{i} \end{pmatrix}, \quad A_3 = \begin{pmatrix} 5 & 8 & 11 \\ 8 & 13 & 18 \\ 11 & 18 & 25 \end{pmatrix},$$

bestimme eine invertierbare Matrix $S_i \in \text{GL}_3(\mathbb{C})$, sodass $S_i^t A_i S_i = \begin{pmatrix} I_k & 0 \\ 0 & 0 \end{pmatrix}$. Welche der Matrizen A_i sind kongruent über \mathbb{C} .

90. Sei V ein endlich dimensionaler reeller Vektorraum. Weiters seien β und β' zwei symmetrische Bilinearformen auf V mit gleicher Signatur. Zeige, dass ein linearer Isomorphismus $\varphi: V \rightarrow V$ existiert, sodass $\beta' = \varphi^* \beta$, d.h. $\beta'(v, w) = \beta(\varphi(v), \varphi(w))$, für alle $v, w, \in V$. *Hinweis: Satz VII.1.35.* Formuliere und beweise die analoge Aussage für Hermitesche Formen

91. Sei V ein n -dimensionaler reeller Vektorraum und $g: V \times V \rightarrow \mathbb{R}$ eine symmetrische Bilinearform mit Signatur (p, q) , wobei $p + q = n$. Weiters sei B eine geordnete Basis von V , sodass $[g]_B = \begin{pmatrix} I_p & \\ & -I_q \end{pmatrix}$. Zeige, dass

$$O_{p,q} := \left\{ A \in M_{n \times n}(\mathbb{R}) \mid A^t \begin{pmatrix} I_p & \\ & -I_q \end{pmatrix} A = \begin{pmatrix} I_p & \\ & -I_q \end{pmatrix} \right\}$$

bezüglich Matrizenmultiplikation eine Gruppe bildet. Zeige weiters, dass

$$O(V, g) \cong O_{p,q}, \quad \varphi \leftrightarrow [\varphi]_B,$$

ein Gruppenisomorphismus ist. *Bemerkung: Die Gruppe $O_{3,1}$ (und auch die Gruppe $O_{1,3}$) wird Lorentz Gruppe genannt, und spielt in der Physik eine zentrale Rolle.*

92. Bestimme die Signatur folgender reeller quadratischer Formen:

$$q_1 = -x^2$$

$$q_2 = -4x^2 - 34y^2 - 12xy$$

$$q_3 = 4x^2 + 13y^2 - 3z^2 + 8xy + 8xz + 26yz$$

$$q_4 = 9x^2 + 26y^2 + z^2 + 4w^2 + 6xy - 6xz + 6xw + 18yz - 18yw - 14zw$$

93. Bestimme die Signatur folgender symmetrischer reeller Matrizen:

$$A_2 = \begin{pmatrix} 4 & 6 \\ 6 & 4 \end{pmatrix}, \quad A_3 = \begin{pmatrix} 4 & 6 & 2 \\ 6 & 13 & 7 \\ 2 & 7 & 9 \end{pmatrix}, \quad A_4 = \begin{pmatrix} 1 & 2 & -1 & 1 \\ 2 & 5 & -1 & 3 \\ -1 & -1 & 1 & -2 \\ 1 & 3 & -2 & -2 \end{pmatrix}$$

Gib jeweils Matrizen $S_n \in GL_n(\mathbb{R})$ an, sodass $S_n^t A_n S_n$ die Gestalt $\begin{pmatrix} I_p & \\ & -I_q \\ & & & 0 \end{pmatrix}$ hat.

94. Zeige, dass für die adjungierte Matrix folgenden Rechenregeln gelten:

(a) $A^{**} = A$, für alle $A \in M_{m \times n}(\mathbb{C})$.

(b) $(A + B)^* = A^* + B^*$, für alle $A, B \in M_{m \times n}(\mathbb{C})$.

(c) $(\lambda A)^* = \bar{\lambda} A^*$, für alle $A \in M_{m \times n}(\mathbb{C})$ und $\lambda \in \mathbb{C}$.

(d) $(AB)^* = B^* A^*$, für alle $A \in M_{m \times n}(\mathbb{C})$ und $B \in M_{n \times l}(\mathbb{C})$.

Zeige, dass die Menge der selbstadjungierten Matrizen einen reellen aber keinen komplexen Teilraum von $M_{n \times n}(\mathbb{C})$ bilden.

95. Sei $h: V \times V \rightarrow \mathbb{C}$ eine Hermitesche Form und $q: V \rightarrow \mathbb{R}$, $q(v) = h(v, v)$, die damit assoziierte quadratische Form. Verifiziere folgende Polarisierungsidentität für $v, w \in V$:

$$h(v, w) = \frac{1}{4}(q(v+w) - q(v-w)) - \mathbf{i} \frac{1}{4}(q(v+iw) - q(v-iw)).$$

96. Zeige, dass die Menge der positiv/negativ (semi)definiten Hermiteschen Formen eine konvexe Teilmenge im reellen Vektorraum aller Hermiteschen Formen bildet. *Hinweis: vgl. Bemerkung VII.1.31.* Formuliere und beweise die analoge Aussage für selbstadjungierte Matrizen.

97. Zeige, dass sich jede Matrix $A \in M_{n \times n}(\mathbb{C})$ auf eindeutige Weise in der Form $A = B + \mathbf{i}C$ schreiben lässt, wobei B und C selbstadjungiert sind. *Hinweis: Betrachte $A + A^*$ und $A - A^*$.*

98. Seien $A \in M_{m \times n}(\mathbb{R})$ und $C \in M_{m \times n}(\mathbb{C})$. Zeige:

(a) $A^t A \geq 0$ und $A^t A > 0$, falls $\text{rank}(A) = n$.

(b) $C^* C \geq 0$ und $C^* C > 0$, falls $\text{rank}(C) = n$.

99. Bezeichne $\mathcal{S}_{p,q} \subseteq M_{2 \times 2}^{\text{sym}}(\mathbb{R})$ die Menge der symmetrischen (2×2) -Matrizen mit Signatur (p, q) . Betrachte den linearen Isomorphismus

$$\phi: M_{2 \times 2}^{\text{sym}}(\mathbb{R}) \xrightarrow{\cong} \mathbb{R}^3, \quad \phi \begin{pmatrix} x & z \\ z & y \end{pmatrix} = \begin{pmatrix} x \\ y \\ z \end{pmatrix},$$

und skizziere die Teilmengen $\phi(\mathcal{S}_{2,0})$, $\phi(\mathcal{S}_{1,1})$, $\phi(\mathcal{S}_{0,2})$ von \mathbb{R}^3 . *Hinweis: Sylvesterkriterium*

100. Bestimme die Signatur der Hermiteschen Matrix

$$\begin{pmatrix} 1 & 2 + \mathbf{i} & -\mathbf{i} \\ 2 - \mathbf{i} & 9 & 1 - 2\mathbf{i} \\ \mathbf{i} & 1 + 2\mathbf{i} & 1 \end{pmatrix}.$$

101 (Cosinussatz). Sei V ein Euklidischer Vektorraum, $v, w \in V$, $v \neq 0 \neq w$ und bezeichne α den Winkel zwischen v und w , d.h. $\cos(\alpha) = \frac{\langle v, w \rangle}{\|v\| \|w\|}$. Zeige

$$\|v - w\|^2 = \|v\|^2 + \|w\|^2 - 2\|v\| \|w\| \cos(\alpha).$$

102. Zeige, dass

$$\|x\|_1 := |x_1| + \cdots + |x_n|$$

eine Norm auf \mathbb{R}^n definiert, für die die Parallelogrammgleichung nicht gilt. Schließe, dass $\|\cdot\|_1$ nicht von einem inneren Produkt auf \mathbb{R}^n induziert wird. *Hinweis: siehe Proposition VII.2.8.*

103. Zeige, dass

$$\langle x, y \rangle = x^t \begin{pmatrix} 1 & 2 & 3 \\ 2 & 8 & 12 \\ 3 & 12 & 27 \end{pmatrix} y.$$

ein inneres Produkt auf \mathbb{R}^3 definiert. Wende das Gram-Schmidt Orthonormalisierungsverfahren auf die Standardbasis von \mathbb{R}^3 an, um eine Orthonormalbasis von \mathbb{R}^3 bezüglich des inneren Produkts oben zu bestimmen.

104. Betrachte den von den Vektoren

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

aufgespannten Teilraum W in \mathbb{R}^4 . Verwende das Gram-Schmidt Orthonormalisierungsverfahren, um eine Orthonormalbasis von W sowie einen Normalektor, $v \in W^\perp$, zu bestimmen.

105. Bestimme die QR-Zerlegung der invertierbaren Matrizen

$$A = \begin{pmatrix} 2 & -2 \\ 2 & 4 \end{pmatrix} \quad \text{und} \quad B = \begin{pmatrix} -2 & -4 & -5 \\ 1 & 3 & 6 \\ 1 & 1 & 2 \end{pmatrix}.$$

106 (Operatornorm). Sei $\varphi: V \rightarrow W$ eine lineare Abbildung zwischen endlich-dimensionalen Euklidischen oder unitären Vektorräumen. Zeige

$$\|\varphi\| := \sup_{0 \neq v \in V} \frac{\|\varphi(v)\|}{\|v\|} = \sup_{v \in V, \|v\|=1} \|\varphi(v)\| < \infty,$$

und verifiziere, dass dies eine Norm auf dem Vektorraum $L(V, W)$ definiert die folgende Eigenschaften besitzt:

- (a) $\|\varphi(v)\| \leq \|\varphi\| \|v\|$, für jedes $v \in V$.
- (b) $\|\psi \circ \varphi\| \leq \|\psi\| \|\varphi\|$, für jede weiter lineare Abbildung $\psi: W \rightarrow U$.
- (c) $\|\varphi^*\| = \|\varphi\|$
- (d) $\|\varphi^* \varphi\| = \|\varphi\|^2$

107 (Kreuzprodukt). Sei a, b ein Orthonormalsystem in \mathbb{R}^3 , d.h. $\|a\| = 1 = \|b\|$ und $\langle a, b \rangle = 0$. Zeige, dass $a, b, a \times b$ eine Orthonormalbasis von \mathbb{R}^3 ist. Schließe daraus, dass die Matrix $A := (a|b|a \times b)$ orthogonal ist. Berechne $\det(A)$.

108. Zeige, dass

$$O_n = \{A \in M_{n \times n}(\mathbb{R}) : A^t A = I_n\}$$

eine kompakte Teilmenge von $M_{n \times n}(\mathbb{R}) = \mathbb{R}^{n^2}$ ist. *Hinweis: Zeige, dass O_n beschränkt und abgeschlossen ist; die Norm $\|A\|^2 = \text{tr}(A^t A)$ auf $M_{n \times n}(\mathbb{R})$ ist dabei hilfreich.* Zeige auch, dass

$$U_n = \{A \in M_{n \times n}(\mathbb{C}) : A^* A = I_n\}$$

eine kompakte Teilmenge von $M_{n \times n}(\mathbb{C}) = \mathbb{C}^{n^2} = \mathbb{R}^{2n^2}$ ist.

109. Zeige $\det(A^*) = \overline{\det(A)}$, für jede Matrix $A \in M_{n \times n}(\mathbb{C})$. Folgere daraus, $|\det(A)| = 1$, für jede unitäre Matrix $A \in U_n$. Zeige, analog, dass $\det(A) \in \{\pm 1\}$, für jede orthogonale Matrix $A \in O_n$.

110 (Legendre-Polynome). Betrachte den Euklidischen Vektorraum der stetigen Funktionen, $C^0([-1, 1], \mathbb{R})$, mit innerem Produkt

$$\langle f, g \rangle = \int_{-1}^1 f(x)g(x)dx.$$

Wende das Gram–Schmidt Orthonormalisierungsverfahren auf die linear unabhängige Teilmenge $1, x, x^2, x^3$ an, und bestimme so eine Orthonormalbasis des davon aufgespannten Teilraums.

111. Betrachte den Vektorraum der stetigen Funktionen, $C^0([0, 2\pi], \mathbb{C})$, mit innerem Produkt

$$\langle f, g \rangle = \frac{1}{2\pi} \int_0^{2\pi} \overline{f(t)}g(t)dt.$$

Zeige, dass die Funktionen e^{int} , $n \in \mathbb{Z}$, ein Orthonormalsystem bilden.

112. Zeige, dass die Menge der reellen bzw. komplexen oberen Dreiecksmatrizen mit positiven Diagonaleinträgen eine Untergruppe von $GL_n(\mathbb{R})$ bzw. $GL_n(\mathbb{C})$ bilden.

113. Seien W und W' zwei komplementäre Teilräume eines endlich dimensionalen reellen Vektorraums V , d.h. $V = W \oplus W'$. Zeige, dass ein inneres Produkt auf V existiert, für das $W^\perp = W'$ gilt. Formuliere und beweise die analoge Aussage für komplexe Vektorräume.

114. Bestimme die Matrix (bezüglich der Standardbasis) der Orthogonalprojektion auf den Teilraum

$$W = \left\langle \begin{pmatrix} 2 \\ 2 \\ 2 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 3 \\ -1 \\ 2 \end{pmatrix} \right\rangle$$

von \mathbb{R}^4 , sowie den Abstand $d(v, W)$ des Punktes $v = \begin{pmatrix} 3 \\ 1 \\ 1 \\ -2 \end{pmatrix}$ zu W .

115. Sei V ein endlich dimensionaler Euklidischer oder unitärer Vektorraum und $p: V \rightarrow V$ ein Projektor, $p^2 = p$. Zeige, dass folgende Aussagen äquivalent sind:

- (a) p ist selbstadjungiert, d.h. $p^* = p$.
- (b) $V = \text{img}(p) \oplus \ker(p)$ ist eine orthogonale Zerlegung, d.h. $\ker(p) = \text{img}(p)^\perp$.
- (c) p ist die Orthogonalprojektion auf $\text{img}(p)$.

116. Sei V ein endlich dimensionaler Euklidischer oder unitärer Vektorraum, bezeichne $\iota: W \rightarrow V$ die Inklusion eines Teilraums und $p: V \rightarrow W$ die Orthogonalprojektion. Zeige $\iota^* = p$ und $p^* = \iota$.

117. Sei V ein endlich dimensionaler unitärer Vektorraum. Zeige, dass

$$\langle v, w \rangle_{\mathbb{R}} := \operatorname{Re}(\langle v, w \rangle)$$

ein inneres Produkt auf dem zugrundeliegenden reellen Vektorraum $V_{\mathbb{R}} = V$ liefert. Für eine lineare Abbildung $\varphi: V \rightarrow V$ bezeichne $\varphi_{\mathbb{R}}: V_{\mathbb{R}} \rightarrow V_{\mathbb{R}}$ die selbe Abbildung, aber als reell lineare Abbildung aufgefasst. Zeige weiters:

- (a) Ist $\varphi: V \rightarrow V$ selbstadjungiert, so ist $\varphi_{\mathbb{R}}: V_{\mathbb{R}} \rightarrow V_{\mathbb{R}}$ symmetrisch.
- (b) Ist $\varphi: V \rightarrow V$ unitär, so ist $\varphi_{\mathbb{R}}: V_{\mathbb{R}} \rightarrow V_{\mathbb{R}}$ orthogonal.

118. Sei V ein endlich dimensionaler Euklidischer oder unitärer Vektorraum und $\varphi: V \rightarrow V$ linear. Zeige:

$$\varphi(W) \subseteq W \quad \Leftrightarrow \quad \varphi^*(W^{\perp}) \subseteq W^{\perp}$$

119. Für jede der folgenden symmetrischen Matrizen A bestimme eine orthogonale Matrix U , sodass $U^{-1}AU = U^tAU$ Diagonalgestalt hat:

$$A_2 = \begin{pmatrix} 9 & -1 \\ -1 & 9 \end{pmatrix}, \quad A_3 = \begin{pmatrix} 2 & -1 & 1 \\ -1 & 2 & 1 \\ 1 & 1 & 2 \end{pmatrix}.$$

120. Für jede der folgenden selbstadjungierten Matrizen A bestimme eine unitäre Matrix U , sodass $U^{-1}AU = U^*AU$ Diagonalgestalt hat:

$$A_2 = \begin{pmatrix} 6 & -2\mathbf{i} \\ 2\mathbf{i} & 9 \end{pmatrix}, \quad A_3 = \begin{pmatrix} 1 & -2\mathbf{i} & 2 \\ 2\mathbf{i} & 1 & 2 \\ 2 & 2 & 1 \end{pmatrix}.$$

121. Zeige, dass die Drehung ρ_E^{θ} in Beispiel VII.4.4 nicht von der Wahl der Orthonormalbasis B abhängt. *Hinweis: Die Gruppe SO_2 ist Abelsch.*

122. Seien a_1 und a_2 zwei linear unabhängige Einheitsvektoren eines endlich dimensionalen Euklidischen Vektorraums, $\|a_1\| = 1 = \|a_2\|$, und bezeichne α den Winkel zwischen a_1 und a_2 . Weiters seien $\sigma_{a_1}, \sigma_{a_2}: V \rightarrow V$ die orthogonalen Spiegelungen an a_1^{\perp} bzw. a_2^{\perp} . Zeige, dass die Komposition $\rho := \sigma_{a_1}\sigma_{a_2}: V \rightarrow V$ mit der Drehung in dem von a_1 und a_2 erzeugten 2-dimensionalen Teilraum $E \subseteq V$ um den Winkel $\theta = 2\alpha$ übereinstimmt, d.h. zeige, $\rho = \rho_E^{\theta}$, mit der Notation in Beispiel VII.4.4. *Hinweis: Die Verdoppelungsformel für den Cosinus lautet: $\cos(2\alpha) = 2\cos^2\alpha - 1$.*

Übungen zu “Lineare Algebra und Geometrie 2”

Wintersemester 2012/13

Stefan Haller

1. Sei $\varphi: V \rightarrow V$ eine selbstadjungierte Abbildung auf einem endlich dimensionalen Euklidischen oder unitären Vektorraum V . Weiters bezeichne λ_{\min} den kleinsten und λ_{\max} den größten Eigenwert von φ . Zeige:

$$\lambda_{\min} \operatorname{id}_V \leq \varphi \leq \lambda_{\max} \operatorname{id}_V.$$

2. Seien $\varphi, \psi: V \rightarrow V$ zwei kommutierende semipositive Abbildungen, d.h. $\varphi \geq 0$, $\psi \geq 0$ und $\varphi\psi = \psi\varphi$. Zeige, dass auch $\varphi\psi \geq 0$ gilt.

3. Bestimme die Quadratwurzeln folgender Matrizen:

$$\begin{pmatrix} 5 & 4 \\ 4 & 5 \end{pmatrix}, \quad \begin{pmatrix} 2 & 2\mathbf{i} \\ -2\mathbf{i} & 2 \end{pmatrix}, \quad \begin{pmatrix} 2 & 3 & 1 \\ 3 & 6 & 3 \\ 1 & 3 & 2 \end{pmatrix}.$$

4. Bestimme Polarzerlegungen folgender Matrizen:

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

5. Bestimme Simulärwertzerlegungen folgender Matrizen:

$$\begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \end{pmatrix}, \quad \begin{pmatrix} 0 & 3 & 0 & 0 & 0 \\ 0 & 0 & 0 & 2 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 5 \\ 0 & 0 & 4 & 0 & 0 \end{pmatrix}.$$

6. Für jedes $t \in \mathbb{R}$ zeige:

$$e^{tA} = \begin{pmatrix} \cosh t & \sinh t \\ \sinh t & \cosh t \end{pmatrix}, \quad \text{wobei} \quad A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

7. Sei $A \in M_{n \times n}(\mathbb{C})$ und $\lambda_1, \dots, \lambda_n$ alle Eigenwerte von A , ihrer Vielfachheit entsprechend oft gelistet. Zeige, dass $e^{\lambda_1}, \dots, e^{\lambda_n}$ die Eigenwerte von e^A sind. *Hinweis: A ist triangulierbar.* SchlieÙe daraus:

$$\det(e^A) = e^{\operatorname{tr}(A)}.$$

Gilt diese Gleichung auch für reelle Matrizen?

8. Sei $\mathbb{K} = \mathbb{R}$ oder $\mathbb{K} = \mathbb{C}$. Für $A, B \in M_{n \times n}(\mathbb{K})$ mit $AB = BA$ zeige:

$$e^{A+B} = e^A e^B.$$

Schließe daraus, dass e^A stets invertierbar ist mit Inverser,

$$(e^A)^{-1} = e^{-A}.$$

Zeige dass, $\mathbb{K} \rightarrow \text{GL}_n(\mathbb{K}), t \mapsto e^{tA}$, ein Gruppenhomomorphismus ist und, dass

$$\frac{d}{dt}e^{tA} = Ae^{tA}$$

gilt. Gib auch zwei (nicht kommutierende) Matrizen an, für die $e^{A+B} \neq e^A e^B$.

9. Sei $\mathbb{K} = \mathbb{R}$ oder $\mathbb{K} = \mathbb{C}$. Zeige

$$(e^A)^* = e^{A^*},$$

für jede Matrix $A \in M_{n \times n}(\mathbb{K})$. Zeige weiters:

- (a) Ist A schiefsymmetrisch, d.h. $A^* = -A$, dann ist e^A unitär bzw. orthogonal. Zeige auch, dass sich jede unitäre Matrix in der Form e^A schreiben lässt, wobei $A^* = -A$. Ist A eindeutig bestimmt? Was lässt sich im reellen Fall sagen?
Hinweis: Unitäre Abbildungen sind normal und daher diagonalisierbar.
- (b) Ist A symmetrisch, d.h. $A^* = A$, dann ist e^A positiv. Zeige auch, dass sich jede positive Matrix in der Form e^A schreiben lässt, wobei $A^* = A$. Ist A eindeutig bestimmt?

10. Bestimme die Pseudoinversen folgender Matrizen:

$$(1 \ 2 \ 3 \ 4), \quad \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad \begin{pmatrix} 1 & 0 \\ 2 & 0 \\ 1 & 1 \end{pmatrix}, \quad \begin{pmatrix} 1 & 2 & 1 \\ 2 & 1 & 2 \\ 3 & 3 & 3 \\ 4 & 5 & 4 \end{pmatrix}$$

11. Sei $A = U\Sigma V^*$ eine Singulärwertzerlegung von $A \in M_{m \times n}(\mathbb{C})$, d.h.

$$U \in U_m, \quad V \in U_n, \quad \Sigma = \begin{pmatrix} \sigma_1 & & & 0 & \dots & 0 \\ & \ddots & & \vdots & & \vdots \\ & & \sigma_k & 0 & \dots & 0 \\ 0 & \dots & 0 & 0 & \dots & 0 \\ \vdots & & \vdots & \vdots & & \vdots \\ 0 & \dots & 0 & 0 & \dots & 0 \end{pmatrix}, \quad \sigma_i > 0.$$

Zeige, $A^+ = V\Sigma^+U^*$. Verwende dies um die Pseudoinverse der Matrix

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

aus Aufgabe 5 zu berechnen.

12. Zeige

$$A^+ = (A^*A)^+A^* = A^*(AA^*)^+$$

für jede reelle oder komplexe Matrix A . *Hinweis: Mit Hilfe der Singulärwertzerlegung, lässt sich dies auf Diagonalmatrizen zurückführen. Alternativ, lässt sich auch nachrechnen, dass $(A^*A)^+A^*$ die vier Eigenschaften hat, die A^+ charakterisieren, und daher mit A^+ übereinstimmen muss.*

13. Sei $A \in M_{m \times n}(\mathbb{C})$ eine Matrix mit Singulärwerten $\sigma_1, \dots, \sigma_k$, ihrer Vielfachheit entsprechend oft gelistet. Betrachte nun die quadratische selbstadjungierte Matrix $B \in M_{(n+m) \times (n+m)}(\mathbb{C})$,

$$B = \begin{pmatrix} 0 & A^* \\ A & 0 \end{pmatrix}.$$

Zeige, dass $\sigma_1, \dots, \sigma_k, -\sigma_1, \dots, -\sigma_k$ alle von Null verschiedenen Eigenwerte von B mit Vielfachheit sind. *Hinweis: Wähle eine Singulärwertzerlegung $A = U\Sigma V^*$, zeige*

$$\begin{pmatrix} 0 & U \\ V & 0 \end{pmatrix} \begin{pmatrix} 0 & A^* \\ A & 0 \end{pmatrix} \begin{pmatrix} 0 & V^* \\ U^* & 0 \end{pmatrix} = \begin{pmatrix} 0 & \Sigma \\ \Sigma^* & 0 \end{pmatrix}$$

und berechne die Eigenwerte dieser Matrix.

14. Für jede Matrix $A \in M_{m \times n}(\mathbb{C})$ betrachte

$$\|A\| := \sup_{0 \neq x \in \mathbb{C}^n} \frac{\|Ax\|}{\|x\|}, \quad \|A\|_2 := (\operatorname{tr}(A^*A))^{1/2}, \quad \|A\|_1 := \operatorname{tr}((A^*A)^{1/2}).$$

Für beliebige unitäre Matrizen $U \in U_m$ und $V \in U_n$ zeige:

$$\|UAV^*\| = \|A\|, \quad \|UAV^*\|_2 = \|A\|_2, \quad \|UAV^*\|_1 = \|A\|_1.$$

Schließe daraus

$$\|A\| = \max\{\sigma_1, \dots, \sigma_k\}, \quad \|A\|_2 = \sqrt{\sigma_1^2 + \dots + \sigma_k^2}, \quad \|A\|_1 = \sigma_1 + \dots + \sigma_k,$$

wobei $\sigma_1, \dots, \sigma_k$ die Singulärwerte von A bezeichnen. Zeige auch

$$\|A^*\| = \|A\|, \quad \|A^*\|_2 = \|A\|_2, \quad \|A^*\|_1 = \|A\|_1$$

sowie

$$\|A\| \leq \|A\|_2 \leq \|A\|_1.$$

15. Zeige, dass $\|A\|_1$ aus dem vorangehenden Beispiel eine Norm auf $M_{m \times n}(\mathbb{C})$ definiert (für $\|A\|$ und $\|A\|_2$ haben wir dies bereits an anderer Stelle verifiziert):

- (a) $\|A\|_1 = 0 \Leftrightarrow A = 0$.
- (b) $\|\lambda A\|_1 = |\lambda| \|A\|_1$, für alle $\lambda \in \mathbb{C}$.
- (c) $\|A + A'\|_1 \leq \|A\|_1 + \|A'\|_1$, wobei $A, A' \in M_{m \times n}(\mathbb{C})$.
- (d) $\|AA'\|_1 \leq \|A\|_1 \|A'\|_1$, wobei $A \in M_{m \times n}(\mathbb{C})$ und $A' \in M_{n \times k}(\mathbb{C})$.
- (e) $|\operatorname{tr}(A)| \leq \|A\|_1$, falls $A \in M_{n \times n}(\mathbb{C})$.
- (f) $\|BA\|_1 \leq \|B\| \|A\|_1$, für alle $B \in M_{l \times m}(\mathbb{C})$.
- (g) $\|AB\|_1 \leq \|A\|_1 \|B\|$, für alle $B \in M_{n \times k}(\mathbb{C})$.
- (h) $\|A\|_1 = \sup_{\|B\| \leq 1} |\operatorname{tr}(BA)|$, wobei das Supremum über alle $B \in M_{n \times m}(\mathbb{C})$ mit $\|B\| \leq 1$ genommen wird.

Hinweis: (a) und (b) sind leicht. Verifiziere dann (e) unter Verwendung einer Singulärwertzerlegung $A = U\Sigma V^$ und $|\operatorname{tr}(V^*U\Sigma)| \leq \operatorname{tr}(\Sigma) = \|\Sigma\|_1$. Anleitung für (f) im Fall $l \geq n$: Sei $A = U\Sigma V^*$ eine Singulärwertzerlegung von A , und $BA = WR$ eine Polarzerlegung von BA , d.h. $W \in M_{l \times n}(\mathbb{C})$, $W^*W = I_n$, $R^* = R \geq 0$.*

SchlieÙe daraus: $\|BA\|_1 = \text{tr}(R) = \text{tr}(V^*W^*BU\Sigma) = \sum_i \langle V^*W^*BU\Sigma e_i, e_i \rangle \leq \|V^*W^*BU\| \|\Sigma\|_1 \leq \|B\| \|\Sigma\|_1 = \|B\| \|A\|_1$. (d) und (g) folgen aus (f) und der vorangehenden Aufgabe. Leite nun (h) her, und verwende diese Darstellung um die Dreiecksungleichung (c) zu zeigen.

16. Gegeben seien reelle Zahlen x_i und y_i wie folgt:

i	1	2	3	4	5
x_i	1.1	1.9	2.8	2.8	4.1
y_i	7.9	6.1	4.5	3.9	1.5

Bestimme ein lineares Polynom, $p(x) = a + bx$, $a, b \in \mathbb{R}$, sodass die Summe der Fehlerquadrate, $\sum_{i=1}^5 |p(x_i) - y_i|^2$, minimal wird. Fertige eine Skizze an!

17. Gegeben seien reelle Zahlen x_i und y_i wie folgt:

i	1	2	3	4	5
x_i	-2	-1	0	1	2
y_i	1.1	-0.2	1.3	3.9	9.6

Bestimme ein quadratisches Polynom, $p(x) = a + bx + cx^2$, $a, b, c \in \mathbb{R}$, sodass die Summe der Fehlerquadrate, $\sum_{i=1}^5 |p(x_i) - y_i|^2$, minimal wird. Fertige eine Skizze an!

18. Zeige, dass

$$G := \left\{ \begin{pmatrix} 1 & 0 \\ b & A \end{pmatrix} : b \in \mathbb{K}^k, A \in \text{GL}_k(\mathbb{K}) \right\}$$

eine Untergruppe von $\text{GL}_{k+1}(\mathbb{K})$ bildet. Zeige weiters, dass

$$G \cong \text{Aff}(\mathbb{K}^k), \quad \begin{pmatrix} 1 & 0 \\ b & A \end{pmatrix} \leftrightarrow (v \mapsto Av + b)$$

ein Gruppenisomorphismus ist.

19. Zeige, dass eine Abbildung $\alpha: V \rightarrow W$ genau dann affin ist, wenn sie folgende Eigenschaft besitzt: Für je zwei Punkte $v_1, v_2 \in V$ und jedes $\lambda \in \mathbb{K}$ gilt

$$\alpha(\lambda v_1 + (1 - \lambda)v_2) = \lambda \alpha(v_1) + (1 - \lambda)\alpha(v_2).$$

Hinweis: Gehe induktiv vor und verwende

$$\lambda_1 v_1 + \dots + \lambda_k v_k = (1 - \lambda_k) \left(\frac{\lambda_1}{1 - \lambda_k} v_1 + \dots + \frac{\lambda_{k-1}}{1 - \lambda_k} v_{k-1} \right) + \lambda_k v_k$$

wobei $\lambda_i \in \mathbb{K}$ und $\sum_{i=1}^k \lambda_i = 1$.

20. Zeige, dass eine Teilmenge A eines Vektorraums V genau dann einen affinen Teilraum bildet, wenn sie folgende Eigenschaft besitzt: Mit je zwei verschiedenen Punkten $a_1 \neq a_2 \in A$ liegt auch die Gerade durch a_1 und a_2 zur Gänze in A .

21. Zeige, dass eine Abbildung $\alpha: V \rightarrow W$ genau dann affin ist, wenn ihr Graph, $G_\alpha := \{(v, \alpha(v)) : v \in V\}$, einen affinen Teilraum von $V \times W$ bildet.

22. Zeige, dass die Punkte

$$a_0 = \begin{pmatrix} 1 \\ 2 \end{pmatrix}, \quad a_1 = \begin{pmatrix} 2 \\ 3 \end{pmatrix}, \quad a_2 = \begin{pmatrix} -1 \\ 3 \end{pmatrix}$$

ein affines Koordinatensystem von \mathbb{R}^2 bilden. Bestimme die affinen und baryzentrischen Koordinaten der Punkte

$$\begin{pmatrix} -4 \\ 9 \end{pmatrix}, \quad \begin{pmatrix} 1 \\ 2 \end{pmatrix}, \quad \begin{pmatrix} 2 \\ 3 \end{pmatrix} \quad \text{und} \quad \begin{pmatrix} -1 \\ 3 \end{pmatrix},$$

vgl. Proposition VIII.1.27. Fertige eine Skizze an!

23 (Strahlensatz). Seien a_0, a_1, a_2 drei affin unabhängige Punkte eines Vektorraums V . Zeige, dass $g_1 := \langle a_0, a_1 \rangle_{\text{aff}}$ und $g_2 := \langle a_0, a_2 \rangle_{\text{aff}}$ zwei Geraden bilden, die sich nur in a_0 schneiden, $g_1 \cap g_2 = \{a_0\}$. Weiters sei $a'_1 \in g_1$ mit $a_0 \neq a'_1 \neq a_1$ und $a'_2 \in g_2$ mit $a_0 \neq a'_2 \neq a_2$. Zeige, dass eindeutig Skalare $\lambda_1, \lambda_2 \in \mathbb{K}$ mit

$$a'_1 = \lambda_1 a_1 + (1 - \lambda_1) a_0 \quad \text{und} \quad a'_2 = \lambda_2 a_2 + (1 - \lambda_2) a_0$$

existieren und, dass $0 \neq \lambda_i \neq 1$ gilt. Zeige, dass $h := \langle a_1, a_2 \rangle_{\text{aff}}$ und $h' := \langle a'_1, a'_2 \rangle_{\text{aff}}$ zwei verschiedene Geraden in V sind, für die

$$h \cap g_1 = \{a_1\}, \quad h \cap g_2 = \{a_2\}, \quad h' \cap g_1 = \{a'_1\} \quad \text{und} \quad h' \cap g_2 = \{a'_2\}$$

gilt. Fertige eine Skizze an! Beweise nun den Strahlensatz:

$$h \cap h' = \emptyset \quad \Leftrightarrow \quad \lambda_1 = \lambda_2.$$

24. Sei V ein endlich dimensionaler Euklidischer Vektorraum und $\sigma: V \rightarrow V$ eine Bewegung, sodass $\sigma^2 = \text{id}_V$. Zeige, dass $A := \{v \in V : \sigma(v) = v\}$ einen nicht leeren affinen Teilraum bildet. Zeige weiters, dass σ mit der Spiegelung an A übereinstimmt, vgl. Beispiel VIII.1.37.

25. Seien A und A' zwei affine Teilmräume eines endlich dimensionalen Euklidischen Vektorraums V . Weiters seien $p, p' \in V$ zwei Punkte, sodass $d(A, p) = d(A', p')$. Zeige, dass eine Bewegung $\alpha: V \rightarrow V$ existiert, sodass $\alpha(A) = A'$ und $\alpha(p) = p'$.

26. Bestimme alle Mittelpunkte der quadratischen Funktionen, $Q_i: \mathbb{R}^3 \rightarrow \mathbb{R}$:

$$Q_1 \begin{pmatrix} x \\ y \\ z \end{pmatrix} = x^2 - y^2 + 6x - 4y + 4$$

$$Q_2 \begin{pmatrix} x \\ y \\ z \end{pmatrix} = x^2 + y^2 - z - 17$$

$$Q_3 \begin{pmatrix} x \\ y \\ z \end{pmatrix} = x^2 + 2y^2 + 3z^2 + 2xy + 4yz + 2x + 6y + 8z + 5$$

27. Sei $Q: V \rightarrow \mathbb{K}$ eine quadratische Funktion und $\alpha \in \text{Aff}(V)$. Zeige, dass α die Menge der Mittelpunkte von $Q \circ \alpha$ bijektiv auf die Menge der Mittelpunkte von Q abbildet. *Hinweis: Zeige $\sigma_{\alpha(m)} = \alpha \circ \sigma_m \circ \alpha^{-1}$, wobei σ_m die Spiegelung am Punkt $m \in V$ bezeichnet.*

28. Fertige eine Liste aller affinen Normalformen von Quadriken in \mathbb{C}^3 an, analog zu Beispiel VIII.2.11.

29. Zu jeder der folgenden Quadriken in \mathbb{C}^2 bestimme einen affinen Isomorphismus $\alpha_i: \mathbb{C}^2 \xrightarrow{\cong} \mathbb{C}^2$, sodass $\alpha_i(E_i)$ Normalform hat, vgl. Beispiel VIII.2.11:

$$E_1 = \left\{ \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{C}^2 : x^2 - 4y^2 + 2x + 8y - 4 = 0 \right\}$$

$$E_2 = \left\{ \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{C}^2 : 2x^2 - 3y^2 - xy + 3x - 2y + 1 = 0 \right\}$$

$$E_3 = \left\{ \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{C}^2 : 4xy = 1 \right\}$$

Analog für folgende Quadrik in \mathbb{C}^3 :

$$E_4 = \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{C}^3 : x^2 - y^2 + 3z^2 + 2xz + 2iyz + 2x + 4iy + 12z + 14 = 0 \right\}.$$

30. Zu jeder der folgenden Quadriken in \mathbb{R}^3 bestimme einen affinen Isomorphismus $\alpha_i: \mathbb{R}^3 \xrightarrow{\cong} \mathbb{R}^3$, sodass $\alpha_i(E_i)$ Normalform hat, vgl. Beispiel VIII.2.16:

$$E_1 = \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3 : 4x^2 + 2y^2 + z^2 + 4xy + 2yz + 4x + 6y + 4z + 1 = 0 \right\}$$

$$E_2 = \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3 : x^2 + 5y^2 + z^2 + 4xy - 2yz - 2x - 2y - 3z + 2 = 0 \right\}$$

$$E_3 = \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3 : x^2 - y^2 - z^2 + 2xz + 2yz + 2x - 2y + 6z - 2 = 0 \right\}$$

31. Seien F und F' zwei Punkte in \mathbb{R}^n und $2f := d(F, F')$. Weiters sei $a > f$. Zeige, dass

$$E := \{x \in \mathbb{R}^n : d(x, F) + d(x, F') = 2a\}$$

eine Quadrik ist. Zeige, dass eine Bewegung $\mathbb{R}^n \rightarrow \mathbb{R}^n$ existiert, die E auf das Rotationsellipsoid,

$$E' = \left\{ x \in \mathbb{R}^n : \left(\frac{x_1}{a}\right)^2 + \left(\frac{x_2}{b}\right)^2 + \cdots + \left(\frac{x_n}{b}\right)^2 = 1 \right\},$$

abbildet, wobei $b = \sqrt{a^2 - f^2}$.

32. Sei A eine affine Hyperebene in \mathbb{R}^n und $F \in \mathbb{R}^n$ ein Punkt, der nicht in A liegt. Zeige, dass

$$E := \{x \in \mathbb{R}^n : d(x, A) = d(x, F)\}$$

eine Quadrik ist. Zeige, dass eine Bewegung $\mathbb{R}^n \rightarrow \mathbb{R}^n$ existiert, die E auf das Rotationsparaboloid,

$$E' = \left\{ x \in \mathbb{R}^n : x_1^2 + \cdots + x_{n-1}^2 = 4fx_n \right\},$$

abbildet, wobei $2f := d(A, F)$. Wie sieht die Situation aus, wenn $F \in A$?

33. Betrachte den Kegel

$$K = \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3 : x^2 + y^2 - z^2 = 0 \right\}$$

und eine affine Isometrie $\varphi: \mathbb{R}^2 \rightarrow \mathbb{R}^3$, $\varphi \begin{pmatrix} x \\ y \end{pmatrix} = A \begin{pmatrix} x \\ y \end{pmatrix} + b$, wobei $A \in M_{3 \times 2}(\mathbb{R})$, $A^*A = I_2$, und $b \in \mathbb{R}^3$. Zeige, dass $\varphi^{-1}(K)$ eine Quadrik ist. Welche Quadriken in \mathbb{R}^2 lassen sich in dieser Form darstellen, für eine geeignete affine Isometrie φ ?

34. Betrachte das einschalige Rotationshyperboloid

$$H = \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3 : \left(\frac{x}{a}\right)^2 + \left(\frac{y}{a}\right)^2 - \left(\frac{z}{b}\right)^2 = 1 \right\},$$

wobei $a, b > 0$. Zeige, dass die Gerade

$$g = \left\{ \begin{pmatrix} a \\ 0 \\ 0 \end{pmatrix} + t \begin{pmatrix} 0 \\ a \\ b \end{pmatrix} : t \in \mathbb{R} \right\}$$

in H liegt, d.h. $g \subseteq H$. Zeige weiters, dass wir durch Rotation der Geraden g um die z -Achse, ganz H erhalten, d.h.

$$H = \bigcup_{\theta \in \mathbb{R}} \rho_\theta(g), \quad \rho_\theta: \mathbb{R}^3 \rightarrow \mathbb{R}^3, \quad \rho_\theta = \begin{pmatrix} \cos \theta & -\sin \theta & 0 \\ \sin \theta & \cos \theta & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Gibt es noch weitere Geraden, die in H liegen?

35. Bestimme eine Bewegung, die die Quadrik

$$E = \left\{ \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2 : x^2 + y^2 + 2xy + \sqrt{2}x + 3\sqrt{2}y - 2 = 0 \right\}$$

auf Normalform bringt.

36. Bestimme eine Bewegung, die die Quadrik

$$E = \left\{ \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2 : 52x^2 + 73y^2 + 72xy + 80x + 190y + 25 = 0 \right\}$$

auf Normalform bringt.

37. Sei $a > b > 0$. Bestimme alle Bewegungen $\alpha: \mathbb{R}^2 \rightarrow \mathbb{R}^2$, die die Ellipse

$$E = \left\{ \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2 : \left(\frac{x}{a}\right)^2 + \left(\frac{y}{b}\right)^2 = 1 \right\}$$

bewahren, $\alpha(E) = E$. *Hinweis: Zeige zunächst, dass jede solche Bewegung linear sein muss.*

38. Betrachte folgende drei Punkte in \mathbb{RP}^3 :

$$P_1 = [0 : 1 : 2 : 3], \quad P_2 = [0 : 1 : 2 : 4], \quad P_3 = [1 : 1 : 1 : 1].$$

Zeige, dass $E := \langle P_1, P_2, P_3 \rangle_{\text{proj}}$ eine projektive Ebene ist und bestimme a_i , sodass

$$E = \left\{ [x_0 : x_1 : x_2 : x_3] \in \mathbb{RP}^3 : a_0x_0 + a_1x_1 + a_2x_2 + a_3x_3 = 0 \right\}.$$

39. Bestimme den Durchschnitt folgender drei projektiver Ebenen in \mathbb{RP}^3 :

$$E_1 = \{[x_0 : x_1 : x_2 : x_3] \in \mathbb{RP}^3 \mid x_0 + x_1 + x_2 + x_3 = 0\}$$

$$E_2 = \{[x_0 : x_1 : x_2 : x_3] \in \mathbb{RP}^3 \mid x_0 - x_1 + x_2 - x_3 = 0\}$$

$$E_3 = \{[x_0 : x_1 : x_2 : x_3] \in \mathbb{RP}^3 \mid x_0 + 2x_1 + 3x_2 + 4x_3 = 0\}$$

40. Für $j = 0, \dots, n$ betrachte die affine Hyperebene

$$A_j := e_j + \langle e_0, \dots, e_{j-1}, e_{j+1}, \dots, e_n \rangle$$

in \mathbb{K}^{n+1} und die damit assoziierte Einbettung

$$\iota_j: \mathbb{K}^n \rightarrow \mathbb{K}P^n, \quad \iota_j(x_1, \dots, x_n) := [x_1 : \dots : x_j : 1 : x_{j+1} : \dots : x_n],$$

wobei $e_j \in \mathbb{K}^{n+1}$ den j -ten Einheitsvektor bezeichnet. Zeige,

$$\mathbb{K}P^n = \bigcup_{j=0}^n \iota_j(\mathbb{K}^n),$$

d.h. die Bilder der Einbettungen ι_j überdecken ganz $\mathbb{K}P^n$.

41. Auf der Sphäre $S^n := \{x \in \mathbb{R}^{n+1} : \|x\| = 1\}$ betrachte die Relation $x \sim y :\Leftrightarrow x = \pm y$. Zeige, dass dies eine Äquivalenzrelation bildet, und dass die Inklusion $S^n \rightarrow \mathbb{R}^{n+1}$ eine wohldefinierte Bijektion $S^n / \sim \xrightarrow{\cong} \mathbb{RP}^n$ induziert. Insbesondere erhalten wir so eine Identifikation $\mathbb{RP}^2 = S^2 / \sim$, d.h. die Punkte der projektiven Ebene \mathbb{RP}^2 können mit Paaren von Antipodalpunkten auf S^2 identifiziert werden. Wie lassen sich die projektiven Geraden in diesem Bild beschreiben?

42. Zeige, dass jede Projektivität $\pi: \mathbb{C}P^n \rightarrow \mathbb{C}P^n$ wenigstens einen Fixpunkt besitzt, d.h. es existiert $P \in \mathbb{C}P^n$ mit $\pi(P) = P$. Gib auch eine Projektivität $\pi: \mathbb{RP}^1 \rightarrow \mathbb{RP}^1$ an, die keinen einzigen Fixpunkt hat.

43. Bestimme die sechs Projektivitäten $\mathbb{K}P^1 \rightarrow \mathbb{K}P^1$ in der Form $x \mapsto \frac{ax+b}{cx+d}$, die die Punkte $0, 1, \infty$ permutieren.

44. Bestimme das Doppelverhältnis folgender Punkte in $\mathbb{K}P^1$:

$$P_0 = [1 : 2], \quad P_1 = [3 : 4], \quad P_2 = [5 : 6], \quad P_3 = [7 : 8].$$

45. Zeige, dass die Punkte

$$P_0 = [1 : 1 : 0], \quad P_1 = [0 : 1 : 1], \quad P_2 = [1 : 2 : 1], \quad P_3 = [1 : 0 : -1]$$

auf einer projektiven Gerade in $\mathbb{K}P^2$ liegen, und bestimme ihr Doppelverhältnis.

46. Für vier Punkte P_0, P_1, P_2, P_3 auf einer projektiven Gerade zeige:

$$DV(P_1, P_0, P_2, P_3) = \frac{1}{DV(P_0, P_1, P_2, P_3)} = DV(P_0, P_1, P_3, P_2)$$

sowie

$$DV(P_2, P_1, P_0, P_3) = 1 - DV(P_0, P_1, P_2, P_3) = DV(P_0, P_3, P_2, P_1)$$

47. Sei $\sigma: \mathbb{C}P^n \rightarrow \mathbb{C}P^n$ eine Projektivität mit $\sigma \circ \sigma = \text{id}_{\mathbb{C}P^n}$ und bezeichne die Menge der Fixpunkte mit $\text{Fix}(\sigma) := \{p \in \mathbb{C}P^n : \sigma(p) = p\}$. Zeige,

$$\text{Fix}(\sigma) = P \cup Q,$$

wobei P und Q zwei projektive Teilräume in $\mathbb{C}P^n$ bezeichnen, für die $P \cap Q = \emptyset$ und $\dim(P) + \dim(Q) = \dim(\mathbb{C}P^n) - 1$ gilt. Zeige weiters, dass eine invariante Hyperebene $H \subseteq \mathbb{C}P^n$ existiert, d.h. $\sigma(H) = H$, sodass die Einschränkung $\sigma|_A: A \rightarrow A$ auf den affinen Teil, $A := \mathbb{C}P^n \setminus H$, eine affine Spiegelung ist.

48. Sei $E \subseteq P(V)$ eine projektive Quadrik und $P \subseteq P(V)$ ein projektiver Teilraum. Zeige, dass $E \cap P$ eine projektive Quadrik in P ist.

49. Für jede der folgenden projektiven Quadriken bestimme eine Projektivität, die diese auf Normalform wie in Satz VIII.4.7 bringt:

(a) $E = \{[x : y : z] \in \mathbb{R}P^2 \mid x^2 + 5y^2 + 4z^2 + 4xy - 2xz = 0\}$

(b) $E = \{[x : y : z] \in \mathbb{R}P^2 \mid xy = 0\}$

(c) $E = \{[x : y : z : w] \in \mathbb{R}P^3 \mid x^2 + 2y^2 + 2z^2 + 2xy + 2yz + 2zw = 0\}$

50. Für jede der folgenden affinen Quadriken $E_0 \subseteq \mathbb{R}^3$ bestimme eine projektive Quadrik in $E \subseteq \mathbb{R}P^3$ mit affinem Teil E_0 , d.h. so, dass $\iota^{-1}(E) = E_0$ gilt, wobei $\iota: \mathbb{R}^3 \rightarrow \mathbb{R}P^3$, $\iota\left(\begin{smallmatrix} x \\ y \\ z \end{smallmatrix}\right) = [1 : x : y : z]$, die Standardembedding bezeichnet:

(a) $E_0 = \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3 \mid x^2 + y^2 + z^2 = 1 \right\}$.

(b) $E_0 = \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3 \mid x^2 + y^2 - z^2 = 1 \right\}$.

(c) $E_0 = \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3 \mid x^2 + y^2 = z \right\}$.

(d) $E_0 = \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3 \mid x^2 = 0 \right\}$. (In diesem Fall ist E nicht eindeutig bestimmt)

Bestimme in jedem Beispiel auch die Menge (Quadrik) aller Fernpunkte von E .

51. Für endlich dimensionale Vektorräume V_1, \dots, V_n zeige

$$\dim(V_1 \times \dots \times V_n) = \dim(V_1) + \dots + \dim(V_n).$$

52. Beweise Lemma IX.1.3, d.h. zeige, dass die direkte Summe $\bigoplus_{i \in I} V_i$ zusammen mit den kanonischen Inklusionen $\iota_i: V_i \rightarrow \bigoplus_{j \in I} V_j$ durch die universelle Eigenschaft, bis auf kanonischen Isomorphismus, eindeutig bestimmt ist.

53. Sei b_i , $i \in I$, eine Basis von V . Zeige, dass die linearen Abbildungen $\mathbb{K} \rightarrow V$, $\iota_i(\lambda) := \lambda b_i$, einen Isomorphismus

$$\bigoplus_{i \in I} \mathbb{K} \xrightarrow{\cong} V$$

induzieren.

54. Seien V_i Vektorräume, $i \in I$. Weiters bezeichnen $\pi_i: \prod_{j \in J} V_j \rightarrow V_i$ die kanonischen Projektionen und $\iota_i: V_i \rightarrow \bigoplus_{j \in I} V_j$ die kanonischen Inklusionen. Zeige, dass eine eindeutig bestimmte lineare Abbildung

$$\phi: \bigoplus_{i \in I} V_i \rightarrow \prod_{i \in I} V_i$$

existiert, sodass $\pi_i \circ \phi \circ \iota_j = \delta_{i,j}$, für alle $i, j \in I$. Zeige, dass ϕ für endliche Indexmengen I ein Isomorphismus ist. Zeige weiters, dass ϕ für unendliche Indexmengen zwar injektiv, aber i.A. nicht surjektiv ist.

55. Beweise Proposition IX.2.7, genauer: Für $\varphi, \varphi_1, \varphi_2 \in L(V, V')$, $\psi, \psi_1, \psi_2 \in L(W, W')$, $\varphi' \in L(V', V'')$, $\psi' \in L(W', W'')$ und $\lambda \in \mathbb{K}$ gilt:

- (a) $\text{id}_V \otimes \text{id}_W = \text{id}_{V \otimes W}$
- (b) $(\varphi' \circ \varphi) \otimes (\psi' \circ \psi) = (\varphi' \otimes \psi') \circ (\varphi \otimes \psi)$
- (c) $(\varphi_1 + \varphi_2) \otimes \psi = \varphi_1 \otimes \psi + \varphi_2 \otimes \psi$
- (d) $\varphi \otimes (\psi_1 + \psi_2) = \varphi \otimes \psi_1 + \varphi \otimes \psi_2$
- (e) $(\lambda\varphi) \otimes \psi = \lambda(\varphi \otimes \psi) = \varphi \otimes (\lambda\psi)$

Dabei bezeichnet $\varphi \otimes \psi: V \otimes W \rightarrow V' \otimes W'$ das Tensorprodukt linearer Abbildungen.

56. Seien V und W zwei Vektorräume. Zeige, dass i.A. nicht jedes Element in $V \otimes W$ elementar, d.h. von der Form $v \otimes w$, ist. *Hinweis: Für einen endlich dimensional Vektorraum V gilt $V^* \otimes V = \text{end}(V)$, und die elementaren Tensoren $\alpha \otimes w \in V^* \otimes V$, wobei $\alpha \in V^*$ und $w \in V$, entsprechen den linearen Abbildungen der Form $V \rightarrow V$, $v \mapsto \alpha(v)w$. Zeige, dass die identische Abbildung nicht von dieser Form ist, falls $\dim(V) \geq 2$.*

57. Sei $\varphi: V \rightarrow V$ und $\psi: W \rightarrow W$ zwei lineare Abbildungen zwischen endlich dimensional \mathbb{K} -Vektorräumen. Zeige

$$\text{tr}(V \otimes W \xrightarrow{\varphi \otimes \psi} V \otimes W) = \text{tr}(\varphi) \text{tr}(\psi)$$

und

$$\det(V \otimes W \xrightarrow{\varphi \otimes \psi} V \otimes W) = \det(\varphi)^{\text{tr}(\psi)} \det(\psi)^{\text{tr}(\varphi)}.$$

Hinweis: Betrachte Basen von V und W , und berechne die Darstellung von $\varphi \otimes \psi$ bezüglich der induzierten Basis von $V \otimes W$.

58. Seien $\varphi: V \rightarrow W$ und $\psi: W \rightarrow U$ zwei lineare Abbildungen zwischen endlich dimensional Vektorräumen über \mathbb{K} . Fassen wir $\varphi \in V^* \otimes W$ und $\psi \in W^* \otimes U$ auf, dann ist $\varphi \otimes \psi \in V^* \otimes W \otimes W^* \otimes U$. Bezeichnet $\text{tr}: W \otimes W^* \rightarrow \mathbb{K}$ die kanonische Kontraktion, dann ist $(\text{id}_{V^*} \otimes \text{tr} \otimes \text{id}_U)(\varphi \otimes \psi) \in V^* \otimes \mathbb{K} \otimes U = V^* \otimes U$. Zeige, dass dieses Element genau der Komposition $\psi \circ \varphi: V \rightarrow U$ entspricht.

59. Sei $\mathbb{K} \subseteq \mathbb{L}$ ein Teilkörper, $\varphi: V \rightarrow W$ eine lineare Abbildung zwischen endlich dimensional \mathbb{K} -Vektorräumen, $B = (b_1, \dots, b_n)$ eine geordnete Basis von V und $C = (c_1, \dots, c_m)$ eine geordnete Basis von W . Weiters bezeichnen

$B_{\mathbb{L}} = (b_1 \otimes 1, \dots, b_n \otimes 1)$ und $C_{\mathbb{L}} = (c_1 \otimes 1, \dots, c_n \otimes 1)$ die entsprechenden Basen der \mathbb{L} -Vektorräume $V \otimes_{\mathbb{K}} \mathbb{L}$ und $W \otimes_{\mathbb{K}} \mathbb{L}$. Für die Matrixdarstellung der \mathbb{L} -linearen Abbildung $\varphi \otimes \text{id}_{\mathbb{L}}: V \otimes_{\mathbb{K}} \mathbb{L} \rightarrow W \otimes_{\mathbb{K}} \mathbb{L}$ zeige:

$$[\varphi \otimes \text{id}_{\mathbb{L}}]_{C_{\mathbb{L}} B_{\mathbb{L}}} = [\varphi]_{CB}.$$

Schließe daraus:

$$\text{tr}_{\mathbb{L}}(\varphi \otimes \text{id}_{\mathbb{L}}) = \text{tr}_{\mathbb{K}}(\varphi) \quad \text{und} \quad \det_{\mathbb{L}}(\varphi \otimes \text{id}_{\mathbb{L}}) = \det_{\mathbb{K}}(\varphi).$$

60. Sei V ein komplexer Vektorraum, bezeichne $V_{\mathbb{R}}$ den zugrundeliegenden reellen Vektorraum und betrachte dessen Komplexifizierung, $V_{\mathbb{R}} \otimes_{\mathbb{R}} \mathbb{C}$. Konstruiere einen kanonischen \mathbb{C} -linearen Isomorphismus

$$V_{\mathbb{R}} \otimes_{\mathbb{R}} \mathbb{C} = V \oplus \bar{V},$$

wobei \bar{V} den komplexen Vektorraum bezeichnet, den wir aus V erhalten indem wir die Skalarmultiplikation umdefinieren.

61. Sei $\mathbb{K} \subseteq \mathbb{L}$ ein Teilkörper und V, W zwei Vektorräume über \mathbb{K} . Konstruiere kanonische Isomorphismen von \mathbb{L} -Vektorräumen:

$$(V \oplus W) \otimes_{\mathbb{K}} \mathbb{L} = (V \otimes_{\mathbb{K}} \mathbb{L}) \oplus (W \otimes_{\mathbb{K}} \mathbb{L})$$

und

$$(V \otimes W) \otimes_{\mathbb{K}} \mathbb{L} = (V \otimes_{\mathbb{K}} \mathbb{L}) \otimes_{\mathbb{L}} (W \otimes_{\mathbb{K}} \mathbb{L}).$$

62. Es bezeichne

$$D := \{A \in M_{n \times n}(\mathbb{K}) \mid \forall i > j : A_{ij} = 0\}$$

die Menge aller oberen Dreiecksmatrizen und

$$D^q := \{A \in M_{n \times n}(\mathbb{K}) \mid \forall i \neq j - q : A_{ij} = 0\}$$

die Teilmenge aller Matrizen, deren nicht-triviale Einträge in der q -ten Nebendiagonale liegen. Zeige, dass $D = \bigoplus_{q=0}^{n-1} D^q$ bezüglich Matrizenmultiplikation eine graduierte Algebra bildet.

63. Zeige $V \otimes V = \Lambda^2 V \oplus S^2 V$, für jeden Vektorraum V .

64. Sei $\varphi: V \rightarrow V$ linear und $\dim(V) = n$. Zeige:

$$\Lambda^n \varphi = \det(\varphi) \text{id}_{\Lambda^n}.$$

Hinweis: in Beispiel IX.5.6 haben wir dies bereits für diagonalisierbare φ gezeigt.

65. Sei $\varphi: V \rightarrow V$ linear und $\dim(V) = n$. Zeige:

$$\det(\varphi + z \text{id}_V) = \sum_{q=0}^n \text{tr}(\Lambda^q \varphi) z^{n-q}.$$

Hinweis: in Beispiel IX.5.7 haben wir dies bereits für diagonalisierbare φ gezeigt.

66. Sei $v \in V$. Zeige, dass $i_v: \Lambda^q V^* \rightarrow \Lambda^{q-1} V^*$,

$$(i_v \alpha)(v_2, \dots, v_q) := \alpha(v, v_2, \dots, v_q),$$

eine graduierte Derivation ist, d.h. es gilt

$$i_v(\alpha \wedge \beta) = (i_v \alpha) \wedge \beta + (-1)^{pq} \alpha \wedge i_v \beta,$$

für alle $\alpha \in \Lambda^p V^*$ und $\beta \in \Lambda^q V^*$.