Analytische Geometrie und Lineare Algebra I

Prof. Dr. Ulrich Stuhler*

Skript zur Vorlesung im Wintersemester 2003/2004

*Georg-August-Universität Göttingen, stuhler@uni-math.gwdg.de

Vorwort

Dies ist eine Ausarbeitung meiner Vorlesung über lineare Algebra, die ich im Wintersemester 2003/2004 in Göttingen gehalten habe. Ich danke den vielen Studierenden, die es durch ihre tatkräftige Arbeit ermöglicht haben, dass diese Ausarbeitung noch im WS 2003/2004 fertig geworden ist, ganz herzlich.

Ein Literaturverzeichnis sowie ein Stichwörterverzeichnis und einige wenige Ergänzungen zum Anhang werden in nächster Zukunft noch folgen und dann einfach ins Netz gestellt werden.

U. Stuhler

Verfasst von

Ulrich Stuhler
Mathematisches Institut der Georg-August-Universität Göttingen
Bunsenstraße 1-3
37073 Göttingen
stuhler@uni-math.gwdg.de

als begleitendes Skript zur von ihm im Wintersemester 2003/2004 am Mathematischen Institut der Georg-August-Universität Göttingen gehaltenen Vorlesung „Analytische Geometrie und Lineare Algebra I“.

Vom Papier in die digitale Form gebracht von seinen Hörern Martin Creutziger, Christian Dickmann, Andreas Dirks, Agnes Dörfelt, Kyle Graehl, Lars Kasper, Stefan Klopottek, Jason Mansour, Henrik Schumacher, Peter Sheldrick, Hendrik Söhnholz und Christian Thiemann mit Hilfe von LaTeX 2ε.

Von der digitalen Form wieder auf Papier gebracht von Stefan Koospal und Jürgen Matthes im Mathematischen Institut.

Inhaltsverzeichnis

1 Der n-dimensionale Raum

Eine kurze Erkundungsfahrt

Bekanntlich beschreibt eine lineare Gleichung etwa der Form

$$x_1 + x_2 = 1$$

eine Gerade in der (x_1, x_2) - Koordinatenebene.

Entsprechend ist die Lösungsmenge zum Beispiel der Gleichung

$$x_1 + x_2 + x_3 = 1$$

im (x_1, x_2, x_3) - Koordinatenraum eine Ebene im Raum, also ein lineares zweidimensionales Gebilde. Viele Probleme aus der Mathematik, den Naturwissenschaften oder den Wirtschaftswissenschaften führen ganz natürlich auf Gleichungen mit mehr als drei Variablen. Betrachten wir zum Beispiel die Gleichung

$$x_1 + x_2 + \ldots + x_{10} = 1$$

Nach den Beispielen oben liegt es doch nahe, diese Gleichung „geometrisch" zu interpretieren. Wir legen Punkte fest durch die zehn Koordinaten $(x_1, x_2, \ldots, x_{10})$. Ein Punkt ist also einfach ein (wie man sagt) 10-Tupel von Koordinaten, die Menge aller solcher 10-Tupel $(x_1, x_2, \ldots, x_{10})$ sollte dann der 10-dimensionale Raum sein, den wir der Kürze halber auch als $\mathbb{R}^{10}$ bezeichnen. Entsprechend hatten wir es eingangs mit dem $\mathbb{R}^2$ (die Ebene) bzw. dem $\mathbb{R}^3$ (der Raum) zu tun. Analog sollte die Lösungsmenge der Gleichung

$$x_1 + x_2 + \ldots + x_{10} = 1$$

ein 9-dimensionales lineares Gebilde in unserem 10-dimensionalen Raum $\mathbb{R}^{10}$ sein. Zumindest ist das plausibel. Man kann etwa die Variablen $x_1, x_2, \ldots, x_9$ ganz frei wählen, dann ist x_{10} eindeutig gegeben durch

$$x_{10} = 1 - (x_1 + x_2 + \ldots + x_9)$$

Aber was soll „9-dimensional" in einem 10-dimensionalen Raum denn nun ganz genau heißen? Dafür brauchen wir eine exakte Definition des Begriffs „Dimension" und das ist einer der wesentlichen Punkte der Vorlesung. Natürlich werden wir dies nicht nacheinander für den $\mathbb{R}^1, \mathbb{R}^2, \ldots, \mathbb{R}^{10}, \mathbb{R}^{11}$ usw. durchführen, sondern wir betrachten gleich allgemein den n-dimensionalen Raum $\mathbb{R}^n$, wo n eine der Zahlen $1, 2, 3, \ldots$ sein kann. Ein Punkt ist dann also einfach ein n-Tupel reeller Zahlen $(x_1, x_2, \ldots, x_n)$. Die Menge aller solcher n-Tupel ist der $\mathbb{R}^n$. Eine einzelne lineare Gleichung mit n Unbekannten hätte jetzt ganz allgemein die Gestalt

$$a_1x_1 + a_2x_2 + \ldots + a_nx_n = b$$

Dabei sind die a_1, a_2, ..., a_n, b feste (aber beliebige) reelle Zahlen, in unserem Beispiel eben war $n = 10$, $a_1 = a_2 = a_3 = \ldots = a_{10} = b = 1$. Die Lösungsmengen derartiger Gleichungen gilt es zu studieren. Statt einer Gleichung tritt auch oft der Fall auf, dass man mehrere Gleichungen simultan zu lösen hat, man spricht dann von einem linearen Gleichungssystem. Das sieht etwa so aus:

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

Dabei sind die a_{11}, a_{12}, a_{21}, ..., a_{mn} sowie die b_1, ..., b_m wieder beliebige aber feste reelle Zahlen. Die x_1, ..., x_n sind die Unbekannten. Man spricht in naheliegender Weise von einem System von m linearen Gleichungen mit n Unbekannten. Die Lösungsmenge eines solchen linearen Gleichungssystems sollte wieder ein „lineares Gebilde“ einer bestimmten Dimension im $\mathbb{R}^n$ sein. Was genau sollte das aber bedeuten? Übrigens werden wir ein derartiges lineares Gleichungssystem bald in viel kompakterer Form schreiben, dann sieht das so aus:

$$\sum_{j=1}^{n} a_{ij}x_j = b_i \qquad (i = 1, \ldots, m)$$

Das spart Platz, aber man muss sich daran gewöhnen! Kann man sich den $\mathbb{R}^n$ nun wirklich anschaulich vorstellen? Hier gilt es, folgendes zu bedenken: Unsere Anschauung ist, wie manchmal scherzhaft gesagt wird, zweieinhalb dimensional. Hierzu ein einfaches Beispiel, das ich einer Vorlesung von H. Grauert entnommen habe. Wir halbieren die Raumdiagonale d eines Würfels im $\mathbb{R}^3$. Im Mittelpunkt P von d errichten wir senkrecht zu d eine Ebene, die mit der Würfeloberfläche zum Schnitt gebracht wird. Welche Figur ergibt sich? Die Figur direkt vor sich zu sehen, ist schwierig. Mit einem einfachen Symmetrieargument folgt aber, dass es sich um ein regelmäßiges Sechseck handeln muss. Das heisst, die Eigenschaften bereits des dreidimensionalen Raumes $\mathbb{R}^3$ sind nur teilweise für uns anschaulich evident. Vieles ergibt sich auch aus logischen Überlegungen sowie aus Analogiebetrachtungen durch Projektion in die Ebene. Ganz ähnlich ergeht es einem nun mit dem n-dimensionalen Raum. Durch Analogie mit dem Raum $\mathbb{R}^3$ werden viele der Aussagen, die wir rein logisch beweisen werden, doch in gewisser Weise einen anschaulichen Gehalt haben. Die Anschauung, in geschickter Weise verwendet, hilft auch oft, die logischen Beweise überhaupt erst zu finden. Es ist genau das Problem einer solchen Vorlesung für die Studierenden, das Ineinandergreifen von logisch völlig strengen Beweisen und leitender Motivation durch anschauliche Vorstellungen zu erfassen.

2 Punkte und Vektoren

Was sind Vektoren

Wie oben beschrieben, ist also ein Punkt P für uns einfach ein Koordinaten-n-Tupel reeller Zahlen $(p_1, \ldots, p_n)$.
Die Menge aller Punkte P ist der n-dimensionale (affine) Raum $\mathbb{R}^n$. Tatsächlich werden zunächst weniger die Punkte des $\mathbb{R}^n$ eine Rolle spielen, sondern viel mehr die Vektoren. Der Grund dafür ist einfach der, dass man mit Vektoren besser als mit Punkten rechnen kann, wie wir gleich sehen werden. Ein Vektor $x = \overrightarrow{PQ}$ wird durch zwei Punkte

$$P = (p_1, p_2, \ldots, p_1)$$

$$Q = (q_1, q_2, \ldots, q_n)$$

des $\mathbb{R}^n$ gegeben. Anschaulich stellen wir uns den Vektor $x = \overrightarrow{PQ}$ als gerichtete Strecke (Vektorpfeil) von P nach Q vor. Sind

$$R = (r_1, r_2, \ldots, r_n)$$

und

$$S = (s_1, s_2, \ldots, s_n)$$

aus dem $\mathbb{R}^n$ zwei weitere Punkte und $y = \overrightarrow{RS}$ der entsprechende Vektor von R nach S, so definieren wir eine Gleichheit(-srelation) zwischen den Vektoren x und y wie folgt:

2.1 Definition: Es ist $x = y$ (bzw. $\overrightarrow{PQ} = \overrightarrow{RS}$) genau dann, wenn gilt:

$$q_i - p_i = s_i - r_i$$

für alle $(i = 1, 2, \ldots, n)$.
Offenbar trifft diese Definition für die „anschaulichen Dimensionen“ $(n = 1, 2, 3)$ genau die geometrische Vorstellung. Für die höheren Dimensionen handelt es sich um eine rein logische Festsetzung, die die anschaulichen Gegebenheiten der niederdimensionalen Räume fortschreibt.
Wir halten folgende Tatsachen fest:

- Ein geordnetes Paar von Punkten $P, Q \in \mathbb{R}^n$ (P, Q aus dem $\mathbb{R}^n$) legt eindeutig einen Vektor $x = \overrightarrow{PQ}$ fest.

- Vektoren $\overrightarrow{PQ}, \overrightarrow{RS}$, sind gleich im Sinne der obigen Definition, wenn gilt

$$q_i - p_i = s_i - r_i \quad (i = 1, \ldots, n)$$

- Vektoren werden also eindeutig durch das so genannte Komponenten-n-Tupel

$$(q_1 - p_1, q_2 - p_2, \ldots, q_n - p_n)$$

beschrieben, d. h. *ebenfalls* durch ein geordnetes n-Tupel reeller Zahlen.

2.2 Bemerkung: Sowohl Punkte als auch Vektoren im $\mathbb{R}^n$ werden also durch n-Tupel reeller Zahlen beschrieben. Trotzdem sind, gemäß ihrer Bedeutung, die beiden Begriffe „Punkt“ und „Vektor“ von einander zu unterscheiden. - Gegeben sei ein Punkt $P = (p_1, \ldots, p_n)$ und ein Vektor $x = (\xi_1, \xi_2, \ldots, \xi_n)$. Dann findet man in

$$Q := (p_1 + \xi_1, p_2 + \xi_2, \ldots, p_n + \xi_n)$$

einen Punkt, so dass gilt $\overrightarrow{PQ}= x$.
Q entsteht aus P durch „Anheften des Vektors x in P“ oder anders gesagt, Q entsteht aus P durch „Verschiebung“ mit dem Vektor x, sind Punkt x. Q ist durch die Bedingung

$$\overrightarrow{PQ}= x$$

eindeutig bestimmt.

Addition von Vektoren

Gegeben seien die Vektoren $x =\overrightarrow{PQ}$ und $y =\overrightarrow{RS}$ durch die Komponenten-n-Tupel

$$\begin{aligned} x &= (q_1 - p_1, \ldots, q_n - p_n) \\ y &= (s_1 - r_1, \ldots, s_n - r_n) \end{aligned}$$

Wir können den Vektor y auch repräsentieren durch $y =\overrightarrow{R'S'}$ mit $R' = Q$ und $S' = (q_1 + (s_1 - r_1), \ldots, q_n + (s_n - r_n)$ Es gilt dann die Gleichheit von Vektoren $\overrightarrow{RS}=\overrightarrow{R'S'}=\overrightarrow{QS'}$. Entsprechend der Addition von Vektoren in den anschaulichen niederdimensionalen Räumen $\mathbb{R}^1, \mathbb{R}^2, \mathbb{R}^3$ setzt man jetzt

$$\overrightarrow{PQ} + \overrightarrow{RS}=\overrightarrow{PQ} + \overrightarrow{R'S'}=\overrightarrow{PQ} + \overrightarrow{QS'}=\overrightarrow{PS'}$$

In Komponenten ergibt sich sofort

$$\overrightarrow{PS'}= ((q_1 - p_1) + (s_1 - r_1), \ldots, (q_n - p_n) + (s_n - r_n))$$

Das legt die folgende Definition nahe, die jetzt nur noch von den Komponenten der Vektoren abhängig und daher unabhängig von der Wahl der Repräsentatoren $x = \overrightarrow{PQ}$ etc. ist.

2.3 Definition: Seien $x = (\xi_1, \ldots, \xi_n),\ y = (\eta_1, \ldots, \eta_n)$ zwei Vektoren im $\mathbb{R}^n$, so definiert man

$$x + y = (\xi_1 + \eta_1, \ldots, \xi_n + \eta_n)$$

Es gelten folgende Regeln für die Addition von Vektoren:

1. Es ist $x + y = y + x$ für $x, y \in \mathbb{R}^n$
   (Kommutativgesetz der Addition)

2. Es ist $(x + y) + z = x + (y + z)$ für $x, y, z \in \mathbb{R}^n$
   (Assoziativgesetz der Addition)

3. Für den Vektor $0 = (0, \ldots, 0)$ (den Nullvektor), der repräsentiert wird durch $\overrightarrow{PP}$ mit beliebigen Punkt $P \in \mathbb{R}^n$, gilt:
   $x + 0 = x$ für alle $x \in \mathbb{R}^n$
   (Existenz eines Nullelements bzw. eines sog. neutralen Elements)

4. Zu zwei Vektoren $a, b \in \mathbb{R}^n$ gibt es genau einen Vektor $x \in \mathbb{R}^n$, der die Gleichung $a + x = b$ löst.

2.4 Bemerkung:

i) Ist $a = (\alpha_1, \ldots, \alpha_n), b = (\beta_1, \ldots, \beta_n)$ so ergibt sich sofort
   $x = (\beta_1 - \alpha_1, \ldots \beta_n - \alpha_n)$.

ii) Ist speziell $b = 0$, so schreiben wir für $x = (-\alpha_1 \ldots, \alpha_n) =: -a$.

iii) Für x in 4. ergibt sich offenbar

$$x = b + (-a)$$

   oder in verkürzter Schreibweise $x = b - a$

Demnach kann man also Vektoren so addieren (auch subtrahieren) wie Zahlen. Wie steht es mit der Multiplikation von Vektoren?

Es gibt, das werden wir im Verlauf dieses Semesters kennen lernen, das Skalarprodukt von Vektoren. Das Produkt ist dabei aber nicht ein Vektor, sondern eine reelle Zahl (ein sog. Skalar). In der Dimension 3 gibt es das sog. Vektorprodukt, das in der Physik eine Rolle spielt. Hierbei handelt es sich aber um eine Spezialität des $\mathbb{R}^3$. Generell kann man aber sogar beweisen, das es auf dem $\mathbb{R}^n$ für $n > 2$ keine zufriedenstellenden, kommutativen Produktbildungen gibt. Was übrig bleibt, ist folgende sehr einfache Bildung.

Sei $\alpha \in \mathbb{R}$ eine reelle Zahl, $x \in \mathbb{R}^n$ ein Vektor, gegeben durch seine Komponenten $x = (\xi_1, \ldots, \xi_n)$.

2.5 Definition: Wir setzen

$$\alpha \cdot x := (\alpha\xi_1, \dots, \alpha\xi_n) \in \mathbb{R}$$

2.6 Bemerkung: Anschaulich bedeutet dies etwa im $\mathbb{R}^3$, das αx wieder in der durch x festgelegten Geraden durch den Nullpunkt liegt. Es gelten die folgenden Rechenregeln für die Multiplikation mit Skalaren:

5. $(\alpha + \beta)x = \alpha x + \beta x$ für $\alpha, \beta \in \mathbb{R}, x \in \mathbb{R}$
6. $\alpha(x + y) = \alpha x + \alpha y$ für $\alpha \in \mathbb{R}, x, y \in \mathbb{R}^n$
7. $\alpha(\beta x) = (\alpha\beta)x$ für $\alpha, \beta \in \mathbb{R}$ und $x \in \mathbb{R}^n$
8. Für das Einselement $1 \in \mathbb{R}$ gilt:
   $1 \cdot x$ für $x \in \mathbb{R}^n$

2.7 Bemerkung: Die Regeln 1–8 beschreiben das, was wir demnächst als reellen Vektorraum festlegen wollen. Alle übrigen Regeln für das Rechnen mit Vektoren können formal durch rein logisches Folgern auf die Regeln 1–8 zurückgeführt werden.

3 Vektorräume

Am Anfang der Vorlesung haben wir den $\mathbb{R}^n$ kennengelernt, ihn als Vektorraum aufgefasst und die grundlegenden Eigenschaften des Rechnens mit Vektoren zusammengefasst. Durch Abstraktion werden wir hieraus unmittelbar den Begriff des abstrakten Vektorraums erhalten.

Als Skalarbereich fixieren wir dabei ein für alle mal einen Körper $(K, +, \cdot)$. Der Leser, insbesondere der an physikalischen Anwendungen interessierte, kann dabei immer an die reellen oder falls gewünscht, komplexen Zahlen denken, bis wir in **Kap. 5a** den abstrakten Körperbegriff einführen. Alternativ kann man natürlich auch zuerst **Kap. 5a** durchlesen.

3.1 Definition: Ein Vektorraum über dem Skalarenkörper K ist eine Menge V (deren Elemente die Vektoren des Vektorraums sind), auf der eine Vektoraddition

$$\begin{aligned} "+" : V \times V &\to V \\ (x, y) &\mapsto x + y \end{aligned}$$

und eine skalare Multiplikation

$$\begin{aligned} "\cdot" : K \times V &\to V \\ (\alpha, x) &\mapsto \alpha \cdot x \end{aligned}$$

erklärt sind, wobei folgende Axiome gelten:

1. "$+$" ist kommutativ, d. h. es gilt: $\quad x + y = y + x \quad$ für alle $x, y \in V$
2. "$+$" ist assoziativ, d. h. es gilt: $\quad (x + y) + z = x + (y + z) \quad$ für alle $x, y, z \in V$
3. Es gibt ein neutrales Element der Addition, $0 \in V$, so dass gilt:
   $x + 0 = x \quad$ für alle $x \in V$
4. Zu jedem Vektor $x \in V$ existiert ein $y \in V$ mit $x + y = 0$.
5. Für $\alpha, \beta \in K$, $x \in V$ gilt: $\quad (\alpha\beta)x = \alpha(\beta x)$
6. (Distributivgesetze) Es gilt:
   - i) $(\alpha + \beta)x = \alpha x + \beta x \quad$ für α, β, x wie oben
   - ii) $\alpha(x + y) = \alpha x + \beta y \quad$ für $\alpha \in K,\ x, y \in V$
7. Für das Einselement $1 \in K$ gilt: $\quad 1 \cdot x = x \quad$ für alle $x \in V$

Bemerkung: Die Bedingungen (1)–(4) besagen, dass $(V, +)$, also die Menge V, versehen mit der Addition als Verknüpfung, eine sogenannte *abelsche* (oder auch kommutative) *Gruppe* ist (benannt nach N. H. Abel, einem der bedeutendsten Mathematiker des 19. Jahrhunderts, der vor allem über die Theorie der algebraischen Gleichungen und sogenannte elliptische Integrale gearbeitet hat).

Wir geben verschiedene Beispiele, die zeigen, wie vielseitig der Begriff des Vektorraums ist.

3.2 Beispiel:

i) Natürlich der $\mathbb{R}^n$ $(n \geq 0)$ mit üblicher Addition und skalarer Multiplikation (dabei $K = \mathbb{R}$).

ii) Ganz entsprechend $V = K^n$, dabei sind die Elemente (Vektoren) aus V einfach die n-Tupel $(\alpha_1, \dots, \alpha_n)$ mit $\alpha_1, \dots, \alpha_n \in K$. Die Addition ist, ganz entsprechend wie bei (i) komponentenweise erklärt durch:

$$(\alpha_1, \dots, \alpha_n) + (\beta_1, \dots, \beta_n) = (\alpha_1 + \beta_1, \dots, \alpha_n + \beta_n)$$

Die skalare Multiplikation durch:

$$\lambda(\alpha_1, \dots, \alpha_n) = (\lambda\alpha_1, \dots, \lambda\alpha_n)$$

mit $\lambda, \alpha_1, \dots, \alpha_n$ sowie $\beta_1, \dots, \beta_n$ aus K.

Die Vektorraumaxiome (1)–(7) folgen sofort aus den entsprechenden Axiomen für den Körper K.

iii) Speziell ist jeder Körper K also ein Vektorraum über sich selber (als Skalarenbereich).

iv) M sei eine beliebige Menge, V ein gegebener Vektorraum.
$\mathrm{Abb}(M; V) = \{f : M \to V \mid f \text{ Abbildung von } M \text{ nach } V\}$, die Menge der Abbildungen von M nach V, „erbt“ dann in natürlicher Weise die Struktur eines K-Vektorraums.

Wir definieren für $f, g \in \mathrm{Abb}(M; V)$ eine *Addition*

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

für $x \in M$.

Die skalare Multiplikation wird so definiert: Sei $\alpha \in K, f \in \mathrm{Abb}(M; V)$, dann setzt man:

$$(\alpha \cdot f)(x) := \alpha \cdot f(x)$$

für $x \in M$.

Dadurch werden $\alpha \cdot f$ bzw. $f + g$ oben definiert. Die Vektorraumaxiome folgen gewissermaßen punktweise (für jedes $x \in M$) aus den entsprechenden Axiomen für den Vektorraum V.

Wir ziehen einige einfache Folgerungen aus den Vektorraumaxiomen.

3.3 Folgerung: Es gibt *genau ein* neutrales Element 0 der Addition in V.

Beweis: Es gibt wenigstens ein neutrales Element $0 \in V$. Angenommen, $0' \in V$ wäre ein weiteres neutrales Element. Dann ist

$$\begin{aligned} &0 + 0' = 0' \text{ , aber auch} \\ &0 + 0' = 0' + 0 = 0 \end{aligned}$$

nach (1) und (3). Daher ist $0 = 0'$. □

3.4 Folgerung: Zu jedem $x \in V$ gibt es *genau ein* $y \in V$ mit $x + y = 0$

Beweis: Nach Axiom (4) gibt es wenigstens ein solches $y \in V$. Angenommen, es gäbe ein weiteres $y' \in V$ mit $x + y' = 0$. Wir berechnen $(x + y) + y' = 0 + y' = y'$, aber andererseits

$$\begin{aligned} (x + y) + y' &= (x + y') + y \quad \text{mit (1) und (2)} \\ &= 0 + y = y \end{aligned}$$

Also $y = y'$, also gibt es höchstens ein Inverses bezüglich der Addition. □

Bemerkung: Wir schreiben für y: $(-x)$.

3.5 Folgerung: Zu Vektoren $a, b \in V$ gibt es genau einen Vektor $x \in V$ mit $a + x = b$.

Beweis: Es ist

$$\begin{aligned} a + x &= b \\ \Leftrightarrow (-a) + (a + x) &= (-a) + b \\ \Leftrightarrow (-a + a) + x &= (-a) + b \\ \Leftrightarrow x &= (-a + b) =: b - a \end{aligned}$$

Also gibt es genau eine eindeutig bestimmte Lösung obiger Gleichung, nämlich $x = (-a) + b$. □

3.6 Folgerung: Es gilt in V: $\quad 0 \cdot a = 0 \quad$ für alle $a \in V$.

Beweis: Offenbar ist

$$\begin{aligned} &0 \cdot a = (0 + 0) \cdot a = 0 \cdot a + 0 \cdot a \text{ , aber auch} \\ &0 \cdot a = 0 \cdot a + 0 \end{aligned}$$

Auf Grund der Eindeutigkeit der Lösungen von Gleichungen der Form $a + x = b$ folgt also: $0 \cdot a = 0$. □

3.7 Folgerung: Es gilt in V für $\lambda \in K$, $a \in V$:

$$(-\lambda)a = -(\lambda a)$$

Beweis: Offenbar ist

$$0 = 0 \cdot a = (\lambda + (-\lambda))a = \lambda a + (-\lambda)a$$

Daher gilt also

$$\begin{aligned} 0 &= \lambda a + (-\lambda)a \text{ , aber auch} \\ 0 &= \lambda a + -(\lambda a) \end{aligned}$$

Wegen der Eindeutigkeit der Lösungen von Gleichungen der Form $a + x = b$ folgt also: $(-\lambda)a = -(\lambda a)$. □

4 Untervektorräume und affine Teilräume

Die Ebene, zum Beispiel der $\mathbb{R}^2$, enthält Punkte und Geraden. Der Raum, zum Beispiel der $\mathbb{R}^3$ enthält Punkte, Geraden und Ebenen. Es liegt nahe zu vermuten, dass diese Begriffe in der Situation abstrakter Vektorräume ein Analogon haben. Da wir vom Begriff des Vektorraumes ausgehen und Vektoren in unserem Aufbau den Vorzug vor Punkten haben, befassen wir uns entsprechend zunächst mit dem mehr algebraischen Begriff eines Untervektorraumes.

4.1 Definition: Sei V ein beliebiger K-Vektorraum. Eine Teilmenge W von V heißt *Untervektorraum* von V, wenn folgendes gilt:

i) Mit zwei Vektoren $x, y \in W$ ist auch $x + y \in W$

ii) Ist $x \in W$, $\lambda \in K$ ein Skalar, so ist $\lambda x \in W$

iii) Es ist der Nullvektor $0 \in W$

4.2 Beispiele:

i) Die Vektoren „auf einer Geraden bzw. Ebene“ durch den Nullpunkt im $\mathbb{R}^3$ bilden einen Untervektorraum des $\mathbb{R}^3$

ii) Sind $x_1,\ldots,x_r \in V$ beliebige Vektoren, so ist die Menge

$$W := \left\{ \sum_{i=1}^{r} \xi_i x_i \mid \xi_1, \ldots, \xi_r \in K \text{ beliebig} \right\}$$

ein Untervektorraum von V. Man nennt W auch den vom System von Vektoren

$$\{x_1, \ldots, x_r\}$$

erzeugten oder auch aufgespannten Untervektorraum.

iii) Gegeben sei eine Gleichung der Form

$$\alpha_1 \xi_1 + \cdots + \alpha_n \xi_n = 0 \quad ,$$

dabei seien $\alpha_1, \ldots, \alpha_n \in K$, die $\xi_1, \ldots, \xi_n$ sind die Unbekannten des Gleichungssystems, so dass eine beliebige Lösung $(\xi_1, \ldots, \xi_n)$ als Vektor des K^n aufgefasst werden kann.

$$W := \{(\xi_1, \ldots, \xi_n) \mid \text{Es ist } \alpha_1\xi_1 + \cdots + \alpha_n\xi_n = 0\}$$

ist ein Untervektorraum des K^n.

Beweis: Übung! □

Man nennt eine Gleichung obiger Form eine homogene lineare Gleichung in den Unbekannten $\xi_1, \ldots, \xi_n$. Ganz entsprechend betrachtet man allgemeiner ein System linearer Gleichungen

$$\sum_{j=1}^{n} \alpha_{ij}\xi_j = 0 \quad (i = 1, \ldots, m)$$

vom m homogenen linearen Gleichungen in den n Unbekannten $\xi_1, \xi_2, \ldots, \xi_n$. Die fest vorgegebenen Koeffizienten $\alpha_{ij} \in K$ werden zusammengefasst in Form einer sogenannten $(m \times n)$-Matrix

$$(\alpha_{ij})_{\substack{1\le i\le m\\ 1\le j\le n}} = \begin{pmatrix} \alpha_{11} & \alpha_{12} & \cdots & \alpha_{1n} \\ \ldots & \ldots & \ldots & \ldots \\ \alpha_{m1} & \alpha_{m2} & \cdots & \alpha_{mn} \end{pmatrix}$$

Man sieht wieder leicht ein: Die Lösungsmenge des homogenen Gleichungssystems

$$W := \left\{(\xi_1, \xi_2, \ldots, \xi_n) \in K^n \mid \text{Es gilt } \sum_{j=1}^{n} \alpha_{ij}\xi_j = 0 \text{ für } i = 1, \ldots, m\right\}$$

ist ein Untervektorraum des K^n.

4.3 Proposition: Ist W ein Untervektorraum des Vektorraumes V, so ist W mit den Verknüpfungen aus V selber ein Vektorraum.

Beweis: Zunächst sollte man sich klar machen, dass etwas zu beweisen ist. Da mit $x, y \in W$ auch $x + y \in W$ gilt, ebenso mit $\alpha \in K$ auch $\alpha x \in W$ ist, so ergeben offenbar die Verknüpfungen auf V auch induzierte Verknüpfungen auf W.

Da $0 \in W$, so gibt es ein Nullelement in W und Axiom 3 der Vektorraumaxiome ist erfüllt. Da mit $x \in W$ auch $(-1)x = -x \in W$ nach (ii) oben, so ist auch Axiom 4 der Vektorraumaxiome erfüllt.

Die übrigen Axiome sind automatisch in W erfüllt, da sie ja in der grösseren Menge V erfüllt sind (Dieser logische Schluss wird sehr oft verwendet und ist sehr bequem: sobald man weiss, dass die Verknüpfungen aus der „Unterstruktur“ nicht herausfallen, folgt zwangsläufig, dass die Axiome für die grössere Verknüpfungsstruktur sich auf die Unterstruktur *vererben*). □

4.4 Proposition:

i) Sind W_1,$W_2 \subset V$ Untervektoräume von V, so ist auch der Durchschnitt $W_1 \cap W_2$ ein Untervektorraum von V.

ii) Es gibt einen kleinsten Untervektorraum W von V, der sowohl W_1 als auch W_2 enthält und den wir mit $(W_1 + W_2)$ bezeichnen.

iii) Die Menge $\{0\}$ ist ein Untervektorraum von V.

Beweis: Wir zeigen nur (ii): Nun enthält jeder Untervektorraum W' mit $W' \supset W_1$, $W' \supset W_2$ alle Summen $w = w_1 + w_2$ mit $w_i \in W_i$ für $i = 1, 2$. Wir definieren daher

$$W := \{w \in V \mid \text{Es gibt } w_1 \in W_1, w_2 \in W_2 \text{ mit } w = w_1 + w_2\}$$

Zeigen wir also, daß die so definierte Menge W ein Untervektorraum von V ist, so ist er auch der kleinste Untervektorraum mit $W \supset W_1, W_2$.

Wir rechnen die Axiome (i), (ii), und (iii) aus **4.1** nach:

i) Seien $x = w_1 + w_2 \in W$ und $y = w_1' + w_2' \in W_2$. Dann ist

$$x + y = (w_1 + w_2) + (w_1' + w_2') = (w_1 + w_1') + (w_2 + w_2{}')$$

Mit w_i,$w_i' \in W_i$ folgt $w_i + w_i' \in W_i$ (i=1,2), da W_i ein Untervektorraum ist. Dann ist aber $x + y \in W$ nach Definition von W.

ii) Mit $x = w_1 + w_2 \in W$ und $\lambda \in K$ ist $\lambda x = \lambda w_1 + \lambda w_2$. Aber $\lambda w_i \in W_i$ (i=1,2), da W_1, W_2 Untervektorräume sind. Also ist $\lambda x \in W$.

iii) Da $0 \in W_1$, $0 \in W_2$ nach Definition eines Untervektorraumes, so folgt $0 = 0{+}0 \in W$ nach Definition von W.

Damit sind alle drei Axiome nachgeprüft: W ist ein Untervektorraum. Nach seiner Definition enthält W die Untervektorräume W_1, W_2. □

4.5 Bemerkung: Einige der Konstruktionen oben lassen sich noch wesentlich allgemeiner fassen, was in den Übungen behandelt wird.

Sei V ein beliebiger K-Vektorraum, I eine beliebige Indexmenge, so dass für jedes Element $i \in I$ ein Untervektorraum W_i von V gegeben ist. Völlig entsprechend zu den obigen Konstruktionen kann man den Durchschnitt als

$$\cap_{i \in I} W_i := \{v \in V \mid v \epsilon W_i \text{ für alle } i \in I\}$$

sowie die Summe $\sum_{i \in I} W_i$ als kleinsten Untervektorraum von V bilden, der alle W_i enthält.

Eine häufig auftretende Situation ist die folgende:
Gegeben seien die beiden Untervektorräume W_1,W_2 von V. Es gelte dabei *zusätzlich*

i) $W_1 \cap W_2 = (0)$

ii) $W_1 + W_2 = V$

(Man sagt auch: W_1 und W_2 sind komplementäre Untervektorräume in V)

4.6 Definition: In obiger Situation sagt man: Der Vektorraum V ist die *direkte Summe* der beiden Unterräume W_1 und W_2. Wir schreiben in dieser Situation $V = W_1 \oplus W_2$.

4.7 Proposition: Ein Vektorraum V ist genau dann die direkte Summe seiner beiden Unterräume W_1,W_2, wenn gilt: Jedes Element $v \in V$ kann eindeutig geschrieben werden als

$$v = w_1 + w_2$$

mit $w_i \in W_i$ (i=1,2).

Beweis: Sei etwa $V = W_1 \oplus W_2$ die direkte Summe seiner beiden Unterräume W_1,W_2. Nach Definition kann dann sicher jedes $v \in V$ als Summe

$$v = w_1 + w_2$$

geschrieben werden. Angenommen, wir hätten zwei derartige Darstellungen

$$v = w_1 + w_2 = w_1' + w_2'$$

mit w_i,$w_i' \in W_i$ (i=1,2). Dann folgt sofort:

$$w_1 - w_1' = w_2' - w_2 =: u$$

Daher folgt $u \in W_1$, $u \in W_2$ und damit $u \in W_1 \cap W_2 = (0)$. Daher ist $u = 0$, weshalb $w_1 = w_1'$, $w_2 = w_2'$ folgt. Genau das war zu zeigen.

Umgekehrt sei die eindeutige Darstellbarkeit wie oben gegeben. Sicher ist dann $V = W_1 + W_2$. Angenommen, es wäre $W_1 \cap W_2 \neq (0)$. Dann gäbe es einen Vektor $w \in W_1 \cap W_2$, $w \neq 0$. Mit $w_1 =: w, w_2 =: -w$ gilt dann

$$0 = 0 + 0 = w + (-w) \quad .$$

Die Darstellung des Nullvektors wäre also im Gegensatz zur Voraussetzung nicht eindeutig. Damit folgt die Behauptung. □

4.8 Definition: Eine Teilmenge A eines K-Vektorraumes V heißt *affiner Teilraum* :⇔ A ist von der Form

$$a + W := \{a + w \mid w \in W\} \quad ,$$

wobei $W \subset V$ ein Untervektorraum von V ist oder $A = \emptyset$, die leere Teilmenge von V.

4.9 Bemerkungen:

i) Ein affiner Teilraum A ist also, wenn er nicht leer ist, einfach ein um den Vektor a „parallel verschobener“ Untervektorraum W.

ii) Der Vektor a ist durch den affinen Teilraum $a + W$ nicht eindeutig bestimmt. Vielmehr gilt $a + W = a' + W$ genau dann, wenn $a - a' \in W$ ist.

4.10 Beispiele:

i) Sei $W = \{0\}$ der Nullvektorraum von V. Dann ist $A = \{a + 0 = a\}$ einfach die Teilmenge von V, die nur aus dem Vektor a besteht. Wir sprechen vom Punkt $\{a\}$, der also durch den Vektor a gegeben ist, davon aber zu unterscheiden ist.

ii) Gegeben seien zwei verschiedene Vektoren $a, b \in V$. Sei $A := a + K(b - a) = \{a + \lambda(b - a) \mid \lambda \in K\}$. Offenbar enthält der affine Teilraum A die Punkte $\{a\}$ und $\{b\}$ wegen $a = a + 0(b - a)$ sowie $b = a + 1(b - a)$. Man spricht von der Geraden durch die Punkte $\{a\}$ und $\{b\}$.

iii) Jeder Untervektorraum W von V ist wegen $W = 0 + W$ auch ein affiner Teilraum.

4.11 Proposition: Eine Teilmenge A von V ist ein affiner Teilraum $\Leftrightarrow$ zu je zwei verschiedenen Punkten $\{a\}, \{b\} \subset A$ ist auch die Gerade durch $\{a\}$ und $\{b\}$ in A enthalten.

Beweis: Siehe Übungen.

4.12 Ergänzung: In **4.2** hatten wir den von r Vektoren $\{x_1, \ldots, x_r\}$ in V aufgespannten Untervektorraum W von V eingeführt. Auch diese Konstruktion kann leicht verallgemeinert werden:
Sei wieder I eine beliebige Indexmenge, für jedes $i \in I$ sei $x_i \in V$ ein Vektor. Der von dem System der $\{x_i\}(i \in I)$ erzeugte (aufgespannte) Untervektorraum von V ist

$$W := \left\{ \sum_{i \in I} \lambda_i x_i \mid \lambda_i \in K \text{ für } i \in I, \text{ fast alle } \lambda_i = 0 \right\} \quad .$$

Hierbei bedeutet „fast alle“: Alle bis auf endlich viele. Wir schreiben für W auch:

$$W = \sum_{i \in I} K x_i$$

Es ist wieder der kleinste Untervektorraum, der alle Vektoren x_i $(i \in I)$ enthält (bzw. alle Untervektorräume $K x_i$).

5 Der Dimensionsbegriff

Wir hatten schon in der Einführung das Problem genannt, den Begriff der Dimension eines Vektorraumes genau festzulegen. Davon handelt dieser Abschnitt.

5.1 Definition: Ein System von r Vektoren $\{x_1, \ldots, x_r\}$ eines Vektorraumes V heißt *linear unabhängig* $:\Leftrightarrow$ Die Gleichung:

$$\sum_{i=1}^{r} \lambda_i x_i = 0 \quad \text{gilt nur, falls} \quad \lambda_1 = \lambda_2 = \ldots = \lambda_r = 0 \quad \text{gilt.}$$

Andernfalls heißt das System $\{x_1, \ldots, x_r\}$ *linear abhängig*.

5.2 Beispiele:

i) Das sogenannte System der Standardvektoren

$$\{e_1 = (1, 0, \ldots, 0), e_2 = (0, 1, 0, \ldots, 0), \ldots, e_n = (0, \ldots, 0, 1)\}$$

im K^n ist *linear unabhängig*. Wäre nämlich

$$\begin{aligned} \sum_{i=1}^{n} \lambda_i e_i &= 0 \qquad \text{im } K^n\text{, so folgt} \\ (\lambda_1, \ldots, \lambda_n) &= (0, \ldots, 0) \quad \text{der Nullvektor, d. h.} \\ \lambda_1 = \lambda_2 = \ldots = \lambda_n &= 0 \end{aligned}$$

ii) Zwei Vektoren im $\mathbb{R}^2$, die nicht auf einer Gerade durch den Nullpunkt liegen, sind linear unabhängig. Drei oder mehr Vektoren im $\mathbb{R}^2$ sind immer linear abhängig. Das werden wir bald bewiesen haben.

iii) Ganz entsprechend sind drei Vektoren im $\mathbb{R}^3$ linear unabhängig, (genau) wenn sie nicht auf einer Ebene durch den Nullpunkt liegen. Vier oder mehr Vektoren im $\mathbb{R}^3$ sind immer linear abhängig. Auch dies wird bald klar sein.

5.3 Folgerung: Ist $\{x_1, \ldots, x_r\}$ ein System von r linear abhängigen Vektoren, so kann *wenigstens einer* der Vektoren $x_1, \ldots, x_r$ durch die anderen dargestellt werden, d. h. es gibt $i \in \{1, \ldots, r\}$ sowie eine Darstellung

$$x_i = \sum_{\substack{j=1 \\ j \neq i}}^{r} \lambda_j x_j \qquad \text{mit } \lambda_j \in K\ .$$

Beweis: Da $\{x_1, \ldots, x_r\}$ linear abhängig sind, besteht eine Gleichung

$$\sum_{j=1}^{r} \mu_j x_j = 0$$

wobei nicht alle μ_j Null sind. Sei etwa $\mu_i \neq 0$. Dann folgt sofort die Gleichung

$$x_i = - \sum_{\left.\begin{smallmatrix} j=1 \\ j\neq i \end{smallmatrix}\right\}}^{r} \frac{\mu_j}{\mu_i} x_j$$

Setzen wir $\lambda_j := -\mu_j/\mu_i$ so folgt die Behauptung. □

5.4 Folgerung: Ist $\{x_1, \ldots, x_r\}$ ein linear unabhängiges System von Vektoren und ist

$$x = \sum_{j=1}^{r} \lambda_i x_i$$

eine Darstellung eines Vektors $x \in V$, so ist diese Darstellung eindeutig.

Beweis: Sei

$$x = \sum_{i=1}^{r} \mu_i x_i$$

eine weitere Darstellung von x als Linearkombination der Vektoren $x_1, \ldots, x_r$. Es folgt

$$0 = \sum_{i=1}^{r} (\lambda_i - \mu_i) x_i$$

Da $\{x_1, \ldots, x_r\}$ ein linear unabhängiges System ist, so folgt

$$\begin{aligned} \lambda_1 - \mu_1 = \ldots = \lambda_r - \mu_r &= 0 \quad \text{d. h.} \\ \lambda_1 = \mu_1, \ldots, \lambda_r &= \mu_r \end{aligned}$$

□

5.5 Definition: Ein System $\{x_1, \ldots, x_r\}$ ist ein Erzeugendensystem des Vektorraumes V genau dann, wenn jeder Vektor $x \in V$ als Linearkombination

$$x = \sum_{i=1}^{r} \lambda_i x_i$$

dargestellt werden kann.

Bemerkung: Es ist *nicht* gefordert, dass die Darstellung eindeutig ist, d. h. dass die $\lambda_1, \ldots, \lambda_r$ eindeutig bestimmt sind.

5.6 Proposition: Das System $\{x_1, \ldots, x_r\}$ ist ein Erzeugendensystem des Vektorraumes V genau dann, wenn V von $\{x_1, \ldots, x_r\}$ aufgespannt wird, beziehungsweise wenn V der kleinste die $x_1, \ldots, x_r$ enthaltende Untervektorraum von V ist.

Beweis: Alles ist schon in **4.2** beziehungsweise **4.12** gezeigt. □

5.7 Definition: Ein System $\{x_1, \ldots, x_n\}$ in V ist eine *Basis* des Vektorraumes V genau dann, wenn das System $\{x_1, \ldots, x_n\}$ linear unabhängig ist, und ein Erzeugendensystem ist.

5.8 Proposition: Ist $\{x_1, \ldots, x_n\}$ eine Basis des Vektorraumes V, so kann jeder Vektor $x \in V$ *eindeutig* als Linearkombination

$$x = \sum_{i=1}^{n} \xi_i x_i$$

dargestellt werden. Die $\xi_1, \ldots, \xi_n$ sind die *Komponenten* von x bezüglich der Basis $\{x_1, \ldots, x_n\}$

5.9 Beispiele:

i) Im K^n hat man als Standardbasis das System der Vektoren

$$\{e_1 = (1, 0, \ldots, 0), e_2 = (0, 1, 0, \ldots, 0), \ldots, e_n = (0, 0, \ldots, 1)\}$$

Natürlich gibt es viele andere Basen im K^n.

ii) Im $\mathbb{R}^2$ ist jedes System von zwei Vektoren $\{x, y\}$, die nicht beide auf einer Geraden liegen, eine Basis (Beweis kommt gleich viel allgemeiner unten). Man formuliere die entsprechende Aussage für den $\mathbb{R}^3$!

5.10 Proposition: Jedes endliche Erzeugendensystem $\{y_1, \ldots, y_s\}$ eines Vektorraumes V besitzt ein Teilsystem, das eine Basis ist.

Beweis: Sei $M = \{i_1, \ldots, i_r\} \subset \{1, \ldots, s\}$ eine minimale Teilmenge mit der Eigenschaft, dass $\{y_{i_1}, \ldots, y_{i_r}\}$ immer noch ein Erzeugendensystem von V ist. Wäre $\{y_{i_1}, \ldots, y_{i_r}\}$ ein linear abhängiges System, so ließe sich einer der Vektoren y_{i_j} durch die verbleibenden Vektoren ausdrücken, das System

$$\{y_{i_1}, \ldots, y_{i_r}\} \setminus \{y_{i_j}\}$$

wäre dann immer noch ein Erzeugendensystem, im Gegensatz zur Annahme der Minimalität. □

Wir kommen jetzt zu einem zentralen Punkt der Vorlesung.

5.11 Satz: (sogenannter *Steinitz'scher Austauschsatz*)
$\{x_1, \ldots, x_r\}$ sei ein System linear unabhängiger Vektoren, $\{y_1, \ldots, y_s\}$ ein Erzeugendensystem des Vektorraumes V. Dann gibt es $\{i_1, \ldots, i_r\} \subset \{1, \ldots, s\}$, so dass das System von Vektoren $\{y'_1, \ldots, y'_s\}$ mit $y'_{i_j} := x_j$ falls $j = 1, \ldots, r$ sowie sonst $y'_j := y_j$ weiterhin Erzeugendensystem von V ist. Speziell folgt die Ungleichung $r \leq s$.

5.12 Bemerkung: Im Erzeugendensystem $\{y_1, \ldots, y_s\}$ werden also r der Vektoren (nämlich $y_{i_1}, y_{i_2}, \ldots, y_{i_r}$) gegen die Vektoren $x_1, \ldots, x_r$ *ausgetauscht*. Man kann dabei nicht von vornherein voraussagen, welche Indizes $i_1, \ldots, i_r$ gebraucht werden.

Beweis (von 5.11): (Durch vollständige Induktion nach r)
Wir zeigen die Behauptung zunächst für $r = 1$. Da $\{y_1, \ldots, y_s\}$ ein Erzeugendensystem von V ist, kann man x_1 als Linearkombination

$$x_1 = \sum_{i=1}^{s} \lambda_i y_i$$

schreiben. Da nach Vorraussetzung das System $\{x_1\}$ linear unabhängig ist, folgt: $x_1 \neq 0$. Daher sind nicht alle λ_i oben Null. Sei $\lambda_{i_1} \neq 0$. Dann tauschen wir y_{i_1} gegen x_1. Wir setzen also

$$\begin{aligned} y'_j &:= y_j \quad , \text{falls } j \neq i_1 \\ y'_{i_1} &:= x_1 \end{aligned}$$

y_{i_1} wird also mit x_1 ausgetauscht. Auf Grund der Relation

$$\begin{aligned} x_1 &= \sum_{i=1}^{s} \lambda_i y_i \\ \text{folgt leicht:} \qquad y_{i_1} &= \lambda_{i_1}^{-1} x_1 - \sum_{\substack{j=1 \\ j \neq i_1}}^{s} \lambda_{i_1}^{-1} \lambda_j y_j \\ &= \lambda_{i_1}^{-1} y'_{i_1} - \sum_{\substack{j=1 \\ j \neq i_1}}^{s} \lambda_{i_1}^{-1} \lambda_j y'_j \\ \text{Daher folgt:} \qquad y_{i_1} &\in \sum_{j=1}^{s} K y'_j \end{aligned}$$

Wegen $y'_k = y_k$ für $k \neq i_1$ folgt außerdem sofort:

$$y_k \in \sum_{j=1}^{s} K y'_j \quad \text{für } k \in \{1, \ldots, s\},\ k \neq i_1.$$

Dann ist aber

$$V = \sum_{j=1}^{s} K y_j \subset \sum_{j=1}^{s} K y'_j \subset V$$

Offenbar gilt daher die Gleichheit

$$\sum_{j=1}^{s} K y'_j = V$$

d. h. $\{y'_1, \ldots, y'_s\}$ ist ein Erzeugendensystem von V.

Induktionsschritt: Sei der Austauschsatz also für den Fall wahr, dass man ein linear unabhängiges System der Ordnung $r-1$ vorliegen hat.
Sei jetzt $\{x_1 \ldots, x_r\}$ ein linear unabhängiges System der Ordnung r. Nach Induktionsvorraussetzung kann man $i_1, \ldots, i_{r-1} \in \{1, \ldots, s\}$ finden, so dass x_j gegen y_{i_j} für $j = 1, \ldots, r-1$ ausgetauscht werden kann und das neue System $y''_1, \ldots, y''_s$ weiterhin Erzeugendensystem ist. Es bleibt übrig, x_r zu vertauschen.
Da $\{y''_1, \ldots, y''_s\}$ ein Erzeugendensystem ist, findet man eine Darstellung

$$x_r = \sum_{j=1}^{s} \beta_j y''_j \quad \text{mit } \beta_j \in K \ .$$

Man findet nun einen Index $i_r := j \notin \{i_1, \ldots, i_{r-1}\}$ mit $\beta_j \neq 0$!
Andernfalls hätte man eine Darstellung der Form

$$x_r = \sum_{j=1}^{r-1} \beta_j x_j \ .$$

Das würde aber der linearen Unabhängigkeit des Systems $\{x_1, \ldots, x_r\}$ widersprechen. Jetzt tauscht man x_r mit y_{i_r} und schließt wie oben. Damit ist der Steinitz'sche Austauschsatz bewiesen. $\square$

5.13 Satz: $\{e_1, \ldots, e_n\}$ sei Basis des Vektorraumes V. Jede andere Basis hat dann ebenfalls n Elemente.

Beweis: Seien also $\{e_1, \ldots, e_n\}$ und $\{f_1, \ldots, f_m\}$ Basen von V. Nach Satz **5.11** folgt, da $\{e_1, \ldots, e_n\}$ insbesondere ein linear unabhängiges System ist, und $\{f_1, \ldots, f_m\}$ als Basis insbesondere ein Erzeugendensystem ist, $n \leq m$. Andersherum schließt man entsprechend $m \leq n$. Daher folgt $n = m$. $\square$

5.14 Definition: Besitzt ein Vektorraum V überhaupt eine (endliche) Basis $\{e_1, \ldots, e_n\}$, so heißt n die Dimension des Vektorraumes V. Man spricht von einem *n-dimensionalen*

Vektorraum. Besitzt V keine (endliche) Basis, so spricht man von einem *unendlichdimensionalen Vektorraum*.

5.15 Bemerkung: Im zweiten Teil der Vorlesung werden wir sehen, dass auch unendlichdimensionale Vektorräume eine Basis haben (natürlich besteht diese dann nicht mehr aus endlich vielen Vektoren). Die Dimension als Zahl wird dann ersetzt durch die sogenannte Mächtigkeit (im Sinne der Mengenlehre) einer solchen Basis. Als solche wird sie wieder eindeutig bestimmt sein.

5.16 Satz: Sei V ein n-dimensionaler Vektorraum. Dann gilt

i) Jedes linear unabhängige System von n Vektoren $\{x_1, \ldots, x_n\}$ ist notwendig ein Erzeugendensystem, also eine Basis.

ii) Jedes Erzeugendensystem $\{x_1, \ldots, x_n\}$ von V ist notwendig auch linear unabhängig, also eine Basis.

Beweis: Sei $\{y_1, \ldots, y_n\}$ eine Basis von V.

Zu i) Nach dem Steinitz'schen Austauschsatz können wir das System der Vektoren $\{x_1, \ldots, x_n\}$ so gegen das System $\{y_1, \ldots, y_n\}$ tauschen, dass das neue System ein Erzeugendensystem bleibt. Dies besagt aber gerade: das System $\{x_1, x_2, \ldots, x_n\}$ ist ein Erzeugendensystem. Dies zeigt (i).

Zu ii) Nach **5.10** können wir aus dem Erzeugendensystem $\{x_1, \ldots, x_n\}$ so Vektoren weglassen, dass das Teilsystem zugleich Erzeugendensystem bleibt und linear unabhängig ist, also eine Basis aus weniger als n Vektoren wäre. Das geht nach Satz **5.13** nicht, (ii) folgt. □

5.17 Folgerung: Jedes linear unabhängige System $\{x_1, \ldots, x_r\}$ eines endlichdimensionalen Vektorraumes kann zu einer Basis ergänzt werden.

Beweis: Sei $\{y_1, \ldots, y_s\}$ eine Basis von V. Anwendung des Steinitz'schen Austauschsatzes ergibt eine Ergänzung $\{y'_1, \ldots, y'_s\}$ des Systems $\{x_1, \ldots, x_r\}$, in dem die x_j passend gegen y_{i_j} getauscht werden, so dass $\{y'_1, \ldots, y'_s\}$ ein Erzeugendensystem ist. Nach Satz **5.13 (ii)** ist es dann notwendig linear unabhängig, also sogar eine Basis. □

5.18 Folgerung: Sei V ein n-dimensionaler Vektorraum.
Dann ist jedes System $\{x_1, \ldots, x_r\}$ von mehr als n Vektoren (also $r > n$) linear abhängig.

Beweis: Sei $\{y_1, \ldots, y_n\}$ eine Basis. Angenommen, $\{x_1, \ldots, x_r\}$ wäre doch linear unabhängig. Dann könnten wir die Vektoren $x_1, x_2, \ldots, x_r$ gegen passende der $y_1, y_2, \ldots, y_n$ austauschen, so dass das neue System ein Erzeugendensystem (sogar eine Basis) bliebe. Dann ist aber $r \leq n$, im Widerspruch zur Vorraussetznug. □

5.19 Folgerung: Sei V ein endlichdimensionaler Teilraum, $U \subset V$ ein Untervektorraum. Dann gilt:

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

Beweis: Sei etwa $\{x_1, \ldots, x_n\}$ eine Basis von V. Wir zeigen noch einmal folgendes einfache Resultat, was eigentlich in **5.11** enthalten ist.

Ist $\{y_1, y_2, \ldots, y_m\}$ ein linear unabhängiges System von Vektoren aus U, das keine Basis ist (also, mit anderen Worten, kein Erzeugendensystem ist), so findet man einen Vektor $y_{m+1} \in U$, der sich nicht durch die Vektoren $y_1, \ldots, y_m$ linear darstellen lässt, für den also gilt:

$$y_{m+1} \notin \sum_{i=1}^{m} K y_i$$

Dann ist aber, wie man sofort sieht, das System $\{y_1, \ldots, y_{m+1}\}$ linear unabhängig. Dieses Verfahren zur Konstruktion eines Systems $\{y_1, \ldots, y_m\}$ werde solange fortgesetzt, bis $\{y_1, \ldots, y_m\}$ ein Erzeugendensystem von U ist oder $m = n$ gilt. Im ersten Fall sind wir fertig, im zweiten Fall mit **5.16 (i)** ebenfalls. $\square$

5A Der Körperbegriff

Der Bereich der Skalare wird in dieser Vorlesung meistens die Menge $\mathbb{R}$ aller reellen Zahlen sein. Wir benötigen nur die einfachsten Rechenregeln für das Rechnen mit reellen Zahlen. Dies legt nahe, diese Regeln zusammenzustellen. Dies wird uns unmittelbar auf den Begriff der sogenannten Körper führen. Wir haben folgende

5A.1 Definition: Ein Körper $(K, +, \cdot)$ (oder einfach K) ist eine Menge K, auf der zwei sogenannte Verknüpfungen "+" (die sogenannte Addition) und "·" (die Multiplikation des Körpers) festgelegt sind, so dass für zwei beliebige Elemente $\alpha, \beta \in K$ ihre Summe $\alpha+\beta \in K$ sowie ihr Produkt $(\alpha \cdot \beta) \in K$ definiert sind und für die folgende Rechenregeln gelten:

Für $\alpha, \beta, \gamma, \ldots \in K$ beliebig gilt

(1) $(\alpha + \beta) + \gamma = \alpha + (\beta + \gamma)$ (Assoziativgesetz der Addition)

(2) $\alpha + \beta = \beta + \alpha$ (Kommutativgesetz der Addition)

(3) Es gibt ein Element $0 \in K$, so dass $\alpha + 0 = 0 + \alpha = \alpha$ gilt (für alle $\alpha \in K$). (Existenz eines sogenannten neutralen Elements für die Addition)

(4) Zu jedem Element $\alpha \in K$ existiert ein Element $\beta \in K$ mit $\alpha + \beta \quad (= \beta + \alpha) = 0$

(5) $(\alpha \cdot \beta) \cdot \gamma = (\alpha \cdot \beta \cdot \gamma)$ (Assoziativgesetz der Multiplikation)

(6) $\alpha \cdot \beta = \beta \cdot \alpha$ (Kommutativgesetz der Multiplikation)

(7) Es gibt ein Element $1 \in K$, das Einselement im Körper, so dass $\alpha \cdot 1 = 1 \cdot \alpha = \alpha$ gilt für alle $\alpha \in K$. (Existenz eines neutralen Elements für die Multiplikation)

(8) Zu jedem Element $\alpha \neq 0$ in K gibt es ein Element $\beta \in K$ mit $\alpha \cdot \beta = \beta \cdot \alpha = 1$ (Existenz eines inversen Elementes für die Multiplikation)

(9) Es ist $1 \neq 0$, d.h. ein Körper K besitzt wenigstens zwei *verschiedene* Elemente.

Die Verbindung zwischen der Verknüpfung der Addition und der der Multiplikation wird durch das sogenannte Distributivgesetz hergestellt.

(10) Es ist $\alpha(\beta + \gamma) = \alpha\beta + \alpha\gamma$ für alle $\alpha, \beta, \gamma \in K$.

5A.2 Beispiele:

i) $K = \mathbb{R}$, der Körper der gewöhnlichen reellen Zahlen mit üblicher Addition und Multiplikation als Verknüpfung.

ii) $K = \mathbb{Q} = \{\frac{r}{s} | r, s \in \mathbb{Z}, s \neq 0\}$, die Menge aller Brüche aus ganzen Zahlen, mit Addition und Multiplikation als Verknüpfung. Man überzeugt sich leicht, dass $(\mathbb{Q}, +, \cdot)$ ein Körper ist, der Körper der rationalen Zahlen, ein fundamentales Objekt der Zahlentheorie.

iii) Wegen Axiom (9) besitzt jeder Körper wenigstens zwei verschiedene Elemente $0, 1$. Gibt es vielleicht einen Körper aus nur zwei Elementen? Offenbar liegt die Addition und Multiplikation im Wesentlichen fest: Sicherlich soll wegen (3) $0+0 = 0, 1+0 = 0+1 = 1$ sein. Wie steht es mit $1+1$, was alleine noch verbleibt. Wegen (4) gibt es zu 1 wenigstens ein Element β mit $1 + \beta = 0$. β ist nicht 0 wegen $1 + 0 = 1 \neq 0$ (nach (9)). Also ist β notwendig 1. Also $1 + 1 = 0$. Damit liegt die Addition vollkommen fest. Nach (7) wissen wir für die Multiplikation $\alpha \cdot 1 = 1 \cdot \alpha = \alpha$. Es bleibt also noch $0 \cdot 0$ festzulegen. Natürlich setzen wir $0 \cdot 0 = 0$ (unten werden wir sehen, dass das ebenfalls zwangsläufig ist). Damit haben wir gezeigt: Wenn es *überhaupt* einen Körper aus genau zwei Elementen gibt, so sind die Addition und Multiplikation wie oben festzulegen. Dies beweist aber noch *nicht*, dass unsere Menge $\{0, 1\} =: \mathbb{F}_2$ mit dieser Addition und Multiplikation wirklich ein Körper ist. Dazu sind die Axiome (1) bis (10) Fall für Fall genau nachzuprüfen. Etwas lästig sind dabei nur die Regeln (1), (5), (10). Tatsächlich klappt alles und wir haben in $\{\mathbb{F}_2, +, \cdot\}$ einen Körper aus zwei Elementen konstruiert.

iv) Noch etwas Exotisches: Verlangt man Axiom (9) nicht, so wird eine Menge $K = \{0\}$ mit $0+0 = 0$ tatsächlich sämtliche anderen Forderungen (1) bis (10) erfüllen. Es ist insbesondere $1 = 0$. Es ist aber zweckmäßig, dies *nicht* als Körper anzuerkennen.

Ehe wir den für uns in dieser Vorlesung besonders interessanten Körper der komplexen Zahlen konstruieren werden, wollen wir eine Reihe einfacher Folgerungen aus den Axiomen (1) bis (10) ziehen.

5A.3 Folgerung: Es gibt genau ein neutrales Element der Addition (Nullelement) und der Multiplikation (Einselement).

Beweis: Seien $0, 0' \in K$ Elemente aus K mit $\alpha + 0 = 0 + \alpha = \alpha$ bzw. $\alpha + 0' = 0' + \alpha = \alpha$ für alle $\alpha \in K$. Einsetzen von $\alpha = 0'$ bzw. $\alpha = 0$ ergibt $0' + 0 = 0'$ sowie $0' + 0 = 0$. Daher folgt $0' = 0$. Der Beweis für das Einselement geht genauso. □

5A.4 Folgerung:

i) Zu $\alpha \in K$ gibt es *genau ein* Element $\beta \in K$ mit $\alpha + \beta = 0$

ii) Ist $\alpha \neq 0$, so gibt es *genau ein* Element $\gamma \in K$ mit $\alpha \cdot \gamma = 1$.

Wir schreiben $\beta = -\alpha$ bzw. $\gamma = \alpha^{-1}$.

Beweis:

i) Angenommen, wir haben $\beta, \beta' \in K$ mit $\alpha + \beta = 0, \alpha + \beta' = 0$
Betrachte dann

$$(\alpha + \beta) + \beta' = \alpha + (\beta + \beta') = \alpha + (\beta' + \beta) = (\alpha + \beta') + \beta$$

Aber

$$\left.\begin{array}{l}(\alpha + \beta) + \beta' = 0 + \beta' = \beta' \\ (\alpha + \beta') + \beta = 0 + \beta = \beta\end{array}\right\}$$

Daher folgt $\beta = \beta'$.

ii) Entsprechend sei bei der Multiplikation $\gamma, \gamma' \in K$ mit $\alpha \cdot \gamma = 1, \alpha \cdot \gamma' = 1$

Betrachte

$$(\alpha \cdot \gamma) \cdot \gamma' = \alpha \cdot (\gamma \cdot \gamma') = \alpha(\gamma'\gamma) = (\alpha\gamma')\gamma$$

Aber

$$\left.\begin{array}{l}(\alpha\gamma)\gamma' = 1 \cdot \gamma' = \gamma' \\ (\alpha\gamma')\gamma = 1 \cdot \gamma = \gamma\end{array}\right\}$$

Daher folgt $\gamma = \gamma'$.

Daher gibt es in beiden Fällen höchstens ein derartiges Element β bzw. γ. Nach 4) bzw. 8) gibt es aber auch ein Element β bzw. γ. Die Behauptung folgt. □

Allgemeiner gilt die

5A.5 Folgerung:

i) Zu $\alpha, \beta \in K$ existiert genau ein Element $x \in K$, nämlich $x = \beta + (-\alpha)$, mit der Eigenschaft

$$\alpha + x = \beta$$

ii) Ist zusätzlich $\alpha \neq 0$, so gibt es genau ein Element $y \in K$, nämlich $y = \alpha^{-1} \cdot \beta$ mit der Eigenschaft

$$\alpha \cdot y = \beta$$

Beweis: Wir zeigen etwa ii), die beiden Fälle sind ganz analog. Wegen $\alpha \neq 0$ existiert jedenfalls $\alpha^{-1} \in K$. Sicherlich ist $\alpha \cdot (\alpha^{-1}\beta) = (\alpha\alpha^{-1})\beta = 1 \cdot \beta = \beta$. Sei umgekehrt $x \in K$ mit $\alpha \cdot x = \beta$. Dann folgt $\alpha^{-1} \cdot (\alpha x) = \alpha^{-1} \cdot \beta$, aber $\alpha^{-1}(\alpha x) = (\alpha^{-1}\alpha)x = 1 \cdot x = x$, daher $x = \alpha^{-1}\beta$, die Behauptung ii) folgt. i) geht entsprechend. □

5A.6 Folgerung: Sei $\alpha \in K$ beliebig. Dann ist $\alpha \cdot 0 = 0 \cdot \alpha = 0$

Beweis: Betrachte $0+0=0$, daher folgt $\alpha \cdot (0+0) = \alpha \cdot 0$. Nach (10) ist aber $\alpha \cdot (0+0) = \alpha \cdot 0 + \alpha \cdot 0$. Also ist die Gleichung $\alpha \cdot 0 + x = \alpha \cdot 0$ lösbar mit $x = 0$, aber auch mit $\alpha \cdot 0$. Die Lösung ist aber nach Folgerung 5A.5,(i) eindeutig. Daher ist $\alpha \cdot 0 = 0$. $\square$

5A.7 Folgerung:

i) Es ist $-(-\alpha) = \alpha$

ii) Es ist $(-\alpha) \cdot \beta = -(\alpha\beta)$

iii) Es ist $(-\alpha)(-\beta) = \alpha \cdot \beta$

Beweis:

i) $-(-\alpha)$ ist nach 5A.4,(i) die Lösung der Gleichung $(-\alpha) + x = 0$. Aber auch $x = \alpha$ ist eine Lösung dieser Gleichung. Auf Grund der Eindeutigkeit folgt $-(-\alpha) = \alpha$

ii) Es ist $\alpha + (-\alpha) = 0$. Daher ist $(\alpha + (-\alpha))\beta = 0 \cdot \beta = 0$, daher $\alpha\beta + (-\alpha)\beta = 0$ nach 5A.1,(10). Die Gleichung $\alpha \cdot \beta + x = 0$ wird eindeutig gelöst durch $x = -(\alpha\beta)$. Damit folgt nach 5A.4,(i) $-(\alpha\beta) = (-\alpha)\beta$.

iii)
$$(-\alpha)(-\beta) = -(\alpha \cdot (-\beta)) = -(-(\alpha\beta)) = \alpha\beta$$
nach i) und ii) $\square$

Bemerkung: Die Körperaxiome implizieren also ganz zwangsläufig die üblichen Vorzeichenregeln. Ganz entsprechend zeigt man etwa die Regeln $(\alpha^{-1})^{-1} = \alpha$ sowie $(\alpha^{-1}\beta)^{-1} = \beta^{-1}\alpha$, sofern $\alpha \neq 0, \beta \neq 0$. Der Vollständigkeit halber beweisen wir noch die üblichen Regeln der „Bruchrechnung“. Für $\alpha^{-1}\beta$ schreiben wir dabei auch $\frac{\beta}{\alpha}$.

5A.8 Folgerung: Es seien $\alpha, \beta, \gamma, \delta \in K$, $\beta, \delta \neq 0$. Dann gilt

i) Es ist $\frac{\alpha}{\beta} = \frac{\alpha\delta}{\beta\delta}$, speziell ist $\beta\delta \neq 0$.

ii) Es ist $\frac{\alpha}{\beta} \pm \frac{\gamma}{\delta} = \frac{\alpha\delta \pm \beta\gamma}{\beta\delta}$

iii) $\frac{\alpha}{\beta} \cdot \frac{\gamma}{\delta} = \frac{\alpha\gamma}{\beta\delta}$

Beweis: Zunächst zeigen wir: $\beta, \delta \neq 0$, dann ist auch $\beta\delta \neq 0$. Dies zeigen wir indirekt. Angenommen also, es wäre $\beta\delta = 0$. Da $\beta \neq 0$ ist, existiert β^{-1}, aber dann ist $0 = \beta^{-1} \cdot 0 = \beta^{-1}(\beta\delta) = 1 \cdot \delta = \delta$. Also $\delta = 0$ im Widerspruch zur Voraussetzung.

Zu i) Es ist

$$\begin{aligned}\frac{\alpha\delta}{\beta\delta} &= \alpha\delta(\beta\delta)^{-1}\\ &= \alpha\delta(\delta^{-1}\beta^{-1})\\ &= \alpha(\delta\delta^{-1})\beta^{-1} = \alpha\beta^{-1} = \frac{\alpha}{\beta}\end{aligned}$$

ii)

$$\begin{aligned}\frac{\alpha}{\beta} + \frac{\gamma}{\delta} &= \frac{\alpha\delta}{\beta\delta} + \frac{\beta\gamma}{\beta\delta} \quad \text{nach i)}\\ &= (\alpha\delta)(\beta\delta)^{-1} + (\beta\gamma)(\beta\delta)^{-1}\\ &= (\alpha\delta + \beta\gamma)(\beta\delta)^{-1} \quad \text{(Distributivgesetz)}\\ &= \frac{\alpha\delta + \beta\gamma}{\beta\delta}\end{aligned}$$

iii)

$$\begin{aligned}\frac{\alpha}{\beta} \cdot \frac{\gamma}{\delta} &= \alpha\beta^{-1}\gamma\delta^{-1}\\ &= \alpha\gamma(\beta\delta)^{-1} \quad \text{nach } \mathbf{5A.7}\text{, Bemerkung}\\ &= \frac{\alpha\gamma}{\beta\delta}\end{aligned}$$

$\square$

Bemerkung: Nach dem Gesagten ist klar, dass die Axiomatik des Körperbegriffs sehr genau das sogenannte „Buchstabenrechnen“ auf den Begriff bringt. Der Vorteil ist, dass wir die einschlägigen Rechenregeln damit für alle Körper gewinnen, z.B. auch für solche andersartigen Beispiele wie $\mathbb{F}_2$.

5B Der Körper der komplexen Zahlen

Dieser ist nächst dem Körper $\mathbb{R}$ der reellen Zahlen der wichtigste Rechenbereich für die Belange der Physik und der Naturwissenschaften. Natürlich spielt er auch in der Mathematik eine zentrale Rolle.

Es ist erstaunlich, wie lange es historisch gedauert hat, bis die Mathematiker die komplexen Zahlen als gleichberechtigt anerkannt haben. Zwar hat zum Beispiel Euler im 18. Jahrhundert viel mit komplexen Zahlen gerechnet, die schöne Formel

$$e^{ix} = \cos x + i \sin x$$

stammt ja von ihm.

Aber erst im 19. Jahrhundert wurde dann ganz systematisch mit komplexen Zahlen gearbeitet und gesehen, dass sie in vieler Hinsicht sich bei weitem besser als die reellen Zahlen verhalten.

Man möchte also den Bereich der reellen Zahlen erweitern und etwa die quadratische Gleichung

$$x^2 = -1$$

lösen können. Man führt daher „formal" die Wurzel $\sqrt{-1} =: i$ als Lösung der obigen Gleichung ein und rechnet dann „wie gewohnt".
Beispielsweise wird man sicherlich mit $a, b, c, d \in \mathbb{R}$ rechnen:

$$(a + b\sqrt{-1}) + (c + d\sqrt{-1}) = a + c + (b + d)\sqrt{-1}$$

sowie

$$\begin{aligned}(a + b\sqrt{-1})(c + d\sqrt{-1}) &= ac + bc\sqrt{-1} + ad\sqrt{-1} + bd(\sqrt{-1})^2 \\ &= (ac - bd) + (bc + ad)\sqrt{-1} \quad .\end{aligned}$$

Speziell ergibt sich die wichtige Formel

$$(a + b\sqrt{-1})(a - b\sqrt{-1}) = a^2 + b^2 \quad .$$

Damit lässt sich leicht das Inverse einer komplexen Zahl berechnen, nämlich für $a + b\sqrt{-1} \neq 0$

$$\begin{aligned}\frac{1}{a + b\sqrt{-1}} &= \frac{a - b\sqrt{-1}}{(a + b\sqrt{-1})(a - b\sqrt{-1})} \\ &= \frac{a - b\sqrt{-1}}{a^2 + b^2} \\ &= \left(\frac{a}{a^2 + b^2} - \frac{b}{a^2 + b^2}\sqrt{-1}\right) \quad .\end{aligned}$$

Also lassen sich Inverse solcher komplexer Zahlen problemlos bestimmen.

Natürlich ist die obige Herleitung etwas unbefriedigend, man führt $\sqrt{-1}$ formal ein, aber was genau soll das heißen? Und was bedeutet „wir rechnen wie gewohnt"?
Wir geben daher noch einmal folgende Konstruktion, die formal unangreifbar ist, die Entstehung aber etwas verdeckt:

Konstruktion: Wir betrachten den reellen Vektorraum $\mathbb{R}^2$ und legen auf ihm die folgenden Rechenverknüpfungen fest.

Addition: Für zwei beliebige Elemente $(a, b), (c, d) \in \mathbb{R}^2$ sei ihre Summe definiert als

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

Bemerkung: Dies ist also einfach die übliche Addition im Vektorraum $\mathbb{R}^2$.

Multiplikation: Diese wird entsprechend zu der obigen Formel festgelegt:

$$(a, b) \cdot (c, d) := (ac - bd, bc + ad)$$

5B.1 Satz: Der $\mathbb{R}^2$, versehen mit der obigen Addition und Multiplikation, bildet einen Körper, den sogenannten Körper der komplexen Zahlen. Bezeichnung: $\mathbb{C}$.

Beweis: Die Rechenregeln für die Addition sind klar, da es sich um den $\mathbb{R}^2$ handelt. Das Nullelement ist dabei $(0, 0) = 0$.
Die Rechenregeln für die Multiplikation können wir nachrechnen, was dem Leser überlassen bleibe. Es sei nur erwähnt, dass das Einselement als $(1, 0)$ gegeben ist. Ein Element $(a, b) \neq (0, 0) = 0$ hat als inverses Element

$$(a, b)^{-1} = \left(\frac{a}{a^2 + b^2}, \frac{-b}{a^2 + b^2} \right) .$$

Wir werden später die Körperaxiome beinahe ganz ohne Rechnung im Rahmen der Matrizenrechnung einsehen. □

Bemerkung:

i) Der Vorteil bei dieser Einführung der komplexen Zahlen ist offenbar, dass jedweder Mystizismus verschwunden ist. Die „Zahl $\sqrt{-1}$" ist hier einfach das Tupel $(0, 1) \in \mathbb{R}^2$ geworden. Man rechnet sofort nach:

$$(0, 1)^2 = (-1, 0) = -1 \quad ,$$

wobei $1 = (1, 0) \in \mathbb{C}$ das Einselement des Körpers ist.

ii) Weiter bietet sich bei dieser Beschreibung unmittelbar die auf Gauß zurückgehende Interpretation der komplexen Zahlen als Vektoren (oder Punkte) in der Ebene $\mathbb{R}^2$ an.

Jetzt können wir auch leicht auf die alte Schreibweise der komplexen Zahlen zurückkommen. Schreiben wir für $a, b \in \mathbb{R}$ für $(a, 0)$ bzw. $(b, 0)$ einfachheitshalber a bzw. b, für $(0, 1)$ einfachheitshalber i, so ergibt sich sofort die Gleichung

$$(a, b) = a + bi \quad .$$

Wir werden von nun an, wenn wir es mit komplexen Zahlen zu tun haben, wieder vorzugsweise diese Darstellung benutzen.

Komplexe Konjugation: Diese bezeichnet die Abbildung $\mathbb{C} \to \mathbb{C}$

$$z = a + bi \mapsto a - bi =: \overline{a + bi} = \overline{z} \quad ,$$

die folgende schöne Eigenschaften hat:

$$\begin{aligned} \overline{z_1 + z_2} &= \overline{z_1} + \overline{z_2} \\ \overline{z_1 \cdot z_2} &= \overline{z_1} \cdot \overline{z_2} \end{aligned}$$

Man sagt auch: Die Abbildung $z \mapsto \overline{z}$ ist ein *Körperautomorphismus*. Die Rechengesetze bleiben bei einem solchen Automorphismus also genau erhalten.

Mit $z = a + bi$ ist offenbar $z\overline{z} = a^2 + b^2$. Die positive reelle Wurzel

$$(z\overline{z})^{\frac{1}{2}} = (a^2 + b^2)^{\frac{1}{2}} =: |z|$$

bezeichnet den Abstand zum Nullpunkt. Es gilt für $z_1, z_2 \in \mathbb{C}$

$$\begin{aligned} |z_1 z_2| &= \left[(z_1 z_2)\left(\overline{z_1 z_2}\right)\right]^{\frac{1}{2}} \\ &= \left(z_1 z_2 \overline{z_1}\,\overline{z_2}\right)^{\frac{1}{2}} \\ &= (z_1 \overline{z_1})^{\frac{1}{2}} (z_2 \overline{z_2})^{\frac{1}{2}} \\ &= |z_1|\,|z_2| \end{aligned}$$

Bemerkung: Nicht allein, dass man jetzt im Körper $\mathbb{C}$ die Gleichung $x^2 = -1$ lösen kann. Die Sache ist noch viel besser. Jede Polynomgleichung der Form

$$x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0 = 0$$

mit komplexen Zahlen $a_0, a_1, \ldots, a_{n-1}$ hat Nullstellen in $\mathbb{C}$ (und zwar mit Vielfachheit gezählt: n Nullstellen).

Dies besagt der sogenannte Fundamentalsatz der Algebra, den wir hier aber nicht beweisen wollen. Dies geschieht vielmehr erst, je nach Geschmack, in der Algebra-Vorlesung algebraisch, oder mit analytischen Hilfsmitteln in der Funktionentheorie-Vorlesung.

6 Ergänzungen zum Dimensionsbegriff

Wir beweisen zunächst die oft nützliche *Dimensionsformel*.

6.1 Satz: V sei ein Vektorraum über dem Körper K. U_1, U_2 seien Untervektorräume. Dann gilt $\dim(U_1) + \dim(U_2) = \dim(U_1 \cap U_2) + \dim(U_1 + U_2)$.

6.2 Bemerkung: Ist dabei wenigstens einer der Summanden auf einer Seite unendlich dimensional, so ist auch wenigstens einer der Terme auf der anderen Seite unendlich dimensional.

Beweis:

I) Seien zunächst U_1 und U_2 endlich dimensional. Dann ist $(U_1 \cap U_2)$ als Teilraum etwa von U_1 ebenfalls endlichdimensional. Sei $S_{12} := \{e_1, \ldots, e_{n_{1,2}}\}$ eine Basis von $(U_1 \cap U_2)$. Wir ergänzen diese nach **5.17** zu einer Basis.

$$S_1 := \{e_1, \ldots, e_{n_{1,2}}, \ldots, e_{n_1}\}$$

von U_1 bzw.

$$S_2 := \{e_1, \ldots, e_{n_{1,2}}, e'_{n_{1,2}+1}, \ldots, e'_{n_2}\}$$

von U_2.

Wir zeigen, dass dann das System

$$S := \{e_1, \ldots, e_{n_{1,2}}, \ldots, e_{n_1}, e'_{n_{1,2}+1}, \ldots, e'_{n_2}\}$$

eine Basis von $(U_1 + U_2)$ ist.

Danach ist

$$\begin{aligned} \dim(U_1 \cap U_2) &= n_{12} \\ \dim(U_1 + U_2) &= (n_1 + n_2 - n_{12}) \end{aligned}$$

Wegen

$$n_1 + n_2 = n_{12} + (n_1 + n_2 - n_{12})$$

folgt die obige Behauptung.

Wir zeigen also, dass S eine Basis von $(U_1 + U_2)$ ist.

Sicherlich gilt für den von der Menge S erzeugten Untervektorraum $\langle S \rangle$:

$$\langle S \rangle \supset U_1, U_2$$

Daher folgt

$$\langle S\rangle \supset U_1 + U_2.$$

Daher ist S ein Erzeugendensystem des Vektorraums $(U_1 + U_2)$.

Bleibt die lineare Unabhängigkeit des Systems S zu zeigen.

Angenommen, man hätte eine nichttriviale Relation

$$\sum_{i=1}^{n_1} \lambda_i e_i + \sum_{i=n_{12}+1}^{n_2} \lambda'_i e'_i = 0$$

Hieraus folgt:

$$\sum_{i=1}^{n_1} \lambda_i e_i = \sum_{i=n_{12}+1}^{n_2} (-\lambda'_i) e'_i =: v$$

Betrachtet man die linke Seite der Gleichung, so folgt $v \in U_1$, betrachtet man die rechte Seite, so folgt $v \in U_2$. Daher folgt $v \in U_1 \cap U_2$.

Es gibt daher $\lambda'_1, \ldots, \lambda'_{n_{1,2}} \in K$, so dass gilt:

$$v = \sum_{i=1}^{n_{1,2}} \lambda'_i e'_i$$

Einsetzen der rechten Seiten ergibt:

$$\sum_{i=n_{1,2}+1}^{n_2} (-\lambda'_i) e'_i = \sum_{i=1}^{n_{1,2}} \lambda'_i e'_i$$

Insgesamt also:

$$\sum_{i=1}^{n_2} \lambda'_i e'_i = 0$$

Da das System $\{e'_1, \ldots, e'_{n_2}\} = S_2$ als Basis von U_2 linear unabhängig ist, folgt:

$$\lambda'_1 = \ldots = \lambda'_{n_2} = 0$$

Dann folgt aber durch Einsetzen sofort:

$$\sum_{i=1}^{n_1} \lambda_i e_i = 0$$

Da $\{e_1, \ldots, e_{n_1}\} = S_1$ eine Basis von U_1 ist, folgt $\lambda_1 = \ldots = \lambda_{n_1} = 0$
Damit ist die lineare Unabhängigkeit von S gezeigt. Die Behauptung folgt.

II) Wir nehmen jetzt an, dass eine der beiden Dimensionen $\dim(U_1)$ oder $\dim(U_2)$ *unendlich* ist.

Dann ist aber $\dim(U_1 + U_2)$ auch unendlich, denn andernfalls wäre $\dim(U_1 + U_2)$ endlich. Damit wären aber auch die beiden Teilräume U_1, U_2 von $(U_1 + U_2)$ endlichdimensional, im Gegensatz zur Annahme. $\square$

6.3 Folgerung: Gilt für zwei Teilräume U_1, U_2 von V

$$\dim(U_1) + \dim(U_2) > \dim(V)$$

so ist $(U_1 \cap U_2) \neq (0)$, d. h. die beiden Untervektorräume haben *nichttrivialen Durchschnitt*

Beweis: Offenbar muß wegen der Dimensionsgleichung $\dim(U_1 \cap U_2) > 0$ sein. Die Behauptung folgt. $\square$

Konstruktive Lösung einer Grundaufgabe

Wir wollen folgendes Problem lösen. Gegeben sei der Standardvektorraum
$V = K^n$, zum Beispiel der $\mathbb{R}^n$, sowie ein System

$$S = \{a_1, a_2, a_3, \dots, a_m\}$$

von m Vektoren

$$a_i = (\alpha_{i1}, \alpha_{i2}, \dots, \alpha_{in}) \text{ für } i = 1 \dots m.$$

Gesucht sind eine Basis und die Dimension des von S erzeugten Untervektorraums $\langle S \rangle$ von V.

Die folgenden drei Prozesse ändern den von S erzeugten Untervektorraum $\langle S \rangle$ nicht:

1) Vertauschung der Vektoren $\{a_1, a_2, \dots, a_m\}$ im System S.

2) Ersetzen eines Vektors a_i durch ein skalares Vielfaches $\lambda_i a_i$ mit $\lambda_i \neq 0$.

3) Ersetzen eines Vektors a_i aus dem System S durch einen Vektor

$$a_i + \lambda_j a_j \text{ mit } j \in \{1, \dots, m\}, j \neq i, \lambda_j \in K.$$

(Die ausführlichen Beweise bleiben den Übungen vorbehalten, sind aber nicht schwer.)

Wir fassen das System der Vektoren $a_i \quad (i = 1, \dots, m)$ jetzt zu einem rechteckigen Schema von Elementen aus K zusammen:

$$\begin{pmatrix} \alpha_{11} & , & \alpha_{12} & , & \dots & , & \alpha_{1n} \\ \alpha_{21} & , & \alpha_{22} & , & \dots & , & \alpha_{2n} \\ \vdots & , & \vdots & , & & , & \vdots \\ \alpha_{m1} & , & \alpha_{m2} & , & \dots & , & \alpha_{mn} \end{pmatrix}$$

Hierbei ist also die sogenannte i-te *Zeile*

$$(\alpha_{i1}, \alpha_{i2}, \dots, \alpha_{in}) = a_i \in K^n$$

der Vektor a_i. Wir haben m Zeilen und n sogenannte *Spalten*. Die j-te Spalte hat also die Gestalt

$$\begin{pmatrix} \alpha_{1j} \\ \alpha_{2j} \\ \vdots \\ \alpha_{mj} \end{pmatrix} (j = 1, \ldots, n)$$

6.4 Definition: Ein solches Schema heisst *Matrix* oder genauer eine $(m \times n)$-Matrix (m Zeilen, n Spalten).

6.5 Bemerkung: Wir schreiben dafür häufig in kürzerer Form:

$$(\alpha_{ij})_{\substack{1 \le i \le m \\ 1 \le j \le n}} =: A$$

Die oben genannten Prozesse (1), (2), (3) können in Termen der Matrix A also auch so beschrieben werden:

1) bedeutet eine Vertauschung von Zeilen der Matrix,

2) bedeutet Multiplikation der i-ten Zeile mit einem Skalar λ_i,

3) schließlich bedeutet Addition des λ_j -fachen der j-ten Zeile zur i-ten Zeile ($i \neq j$!).

(wobei bei (2), (3) die übrigen Zeilen nicht verändert werden, in (3) wird also insbesondere die j-te Zeile nicht verändert).

Mittels (1), (2) und (3) werden wir jetzt durch Umformung der Matrix eine Basis erhalten: Wir gehen induktiv vor.

Sei i_1 minimal, so dass die i_1-te Spalte der Matrix A eine Komponente $\neq 0$ enthält. Mittels (1) (Vertauschung von Zeilen) können wir annehmen, dass dies gerade die Komponenten α_{1,i_1} ist.

Mittels (2) multiplizieren wir die erste Zeile mit α_{1,i_1}^{-1} und können daher o.E.d.A. $\alpha_{1,i_1} = 1$ annehmen. Schließlich subtrahieren wir jetzt das α_{j,i_1}-fache der ersten Zeile von der j-ten Zeile für $j = 2, \ldots, m$ und können dann o.E.d.A. $\alpha_{j,i_1} = 0$ für $j = 2, \ldots, m$ annehmen.

Induktionsschritt: Sei jetzt die Matrix A bereits so weit bearbeitet, dass

$$\alpha_{1,i_1} = \alpha_{2,i_2} = \ldots = \alpha_{r,i_r} = 1$$

mit $1 \le i_1 < i_2 < \ldots < i_r \le n$ ist (und natürlich $r \le m$), ferner seien die Komponenten α_{jk} der j-ten Zeile für $1 \le j \le r$, die links von α_{j,i_j} stehen (für die also $1 \le k < i_j$ ist) alle Null. Weiter sind die α_{jk} mit $j > r$, $k \le i_r$ alle Null (der Leser mache sich eine Skizze).

Wir betrachten jetzt das System der verbleibenden Zeilenvektoren

$$\{a_{r+1}, \ldots, a_m\}$$

und die sich daraus ergebende Matrix A' (die also aus A durch Streichen der ersten r Zeilen entsteht).

Auf A' wenden wir jetzt den oben beschriebenen ersten Schritt des Verfahrens an und erhalten mit Hilfe der Prozesse (1), (2), (3)

$$\alpha_{1,i_1} = \ldots = \alpha_{r,i_r} = \alpha_{r+1,i_{r+1}} = 1$$

etc. Dabei ist nach Annahme im Induktionsschritt $i_{r+1} > i_r$.

Damit ist der Induktionsschritt vollzogen.

Da $1 \leq i_1 < i_2 < \ldots < i_r < \ldots \leq n$ gelten muß, ist gesichert, dass das Verfahren abbrechen muß, etwa nach dem k-ten Schritt. Die hier eben betrachtete Matrix A' wird dann die Null-Matrix sein, dass sich ergebende System der Vektoren $\{a'_1, \ldots, a'_k\}$ ist sicher ein Erzeugendensystem des Untervektorraumes $\langle a_1, a_2, \ldots, a_m \rangle$ des K^n, der von $a_1, a_2, \ldots, a_m$ aufgespannt wird.

Die lineare Unabhängigkeit sieht man sofort so ein: Angenommen, es bestünde eine Relation

$$\sum_{i=1}^{k} \lambda_i a'_i = 0$$

Wir betrachten die i_1-te Komponente der Vektoren a'_i. Nach Konstruktion gilt

$$\alpha'_{1,i_1} = 1, \alpha'_{2,i_1} = \ldots = \alpha'_{k,i_1} = 0$$

Daher folgt sofort $\lambda_1 = 0$. Entsprechende Betrachtung der $i_2, \ldots, i_k$-ten Komponente liefert die Behauptung. □

Zum Abschluß dieses Abschnitts definieren wir noch, was die Dimension eines affinen Teilraumes eines Vektorraumes V sein soll.

6.6 Definition:

i) Sei $X = a + U$ ein affiner Teilraum von V, wobei also $a \in V$ ein Vektor, $U \subset V$ ein Untervektorraum von V ist. Es sei

$$\dim(X) := \dim(U)$$

ii) Ist $X = \emptyset$ die leere Menge, so sei $\dim(X)$ (per Konvention!) gleich (-1).

6.7 Bemerkung:

i) Die Dimension eines affinen Raumes ist also, falls er nicht leer ist, gegeben durch die Dimension des zu Grunde liegenden Untervektorraumes.

ii) Ist $\dim(X) = \dim(V) - 1$ so spricht man von einer affinen Hyperebene.

7 Lineare Abbildungen

Aus dem Geometrieunterricht in der Schule wird der Leser bereits Beispiele von linearen Abbildungen kennen, etwa Drehungen, zentrische Streckungen, Spiegelungen oder Scherungen. Parallelverschiebungen (Translationen), die ebenfalls eine wichtige Rolle spielen, werden später behandelt, da wir zunächst immer voraussetzen wollen, dass ein Punkt, der Nullvektor, fest bleibt.

Was ist den eben genannten Abbildungen einer Ebene in sich gemeinsam, was man zu einer allgemeinen Definition heranziehen könnte? Offenbar ist das Bild eines Parallelogramms, auch wenn sich Winkel und Seitenlängen verändert haben, wieder ein Parallelogramm. Dies deutet darauf hin, dass das Bild einer Summe von Vektoren die Summe der Bilder der Vektoren sein sollte. Dies führt auf die folgende ganz allgemeine

7.1 Definition: V, W seien Vektorräume über dem Körper K.
Eine Abbildung $\varphi : V \to W$ heißt lineare Abbildung, falls gilt:

i) $\varphi(x+y) = \varphi(x) + \varphi(y)$ für alle Vektoren $x, y \in V$

ii) $\varphi(\lambda x) = \lambda\varphi(x)$ für alle $x \in V, \lambda \in K$

7.2 Folgerung: Für $x_1, \ldots, x_r \in V$, $\lambda_1, \ldots, \lambda_k \in K$ beliebig gilt:

$$\varphi\left(\sum_{i=1}^{r} \lambda_i x_i\right) = \sum_{i=1}^{r} \lambda_i \varphi(x_i)$$

7.3 Bemerkung: Aus ii) folgt sofort $\varphi(0) = \varphi(0 \cdot 0) = 0 \cdot \varphi(0) = 0$ (0 in zwei Bedeutungen, als Element von V bzw. von K).

7.4 Folgerung und Definition: Sei $\varphi : V \to W$ eine lineare Abbildung. $U \subset V$ sei ein Untervektorraum von V. Dann ist die Bildmenge $\varphi(U)$ in W ein Untervektorraum von W. Ist speziell $U = V$, so ist die Menge $\varphi(V)$ als Untervektorraum von W das sog. Bild $\mathrm{Im}(\varphi)$.

Beweis: Seien $\varphi(x), \varphi(y) \in \varphi(U)$ zwei beliebige Vektoren. Dann ist aber ihre Summe $\varphi(x) + \varphi(y) = \varphi(x+y) \in \varphi(U)$ ebenfalls aus $\varphi(U)$. Entsprechend ist $\lambda\varphi(x) = \varphi(\lambda x) \in \varphi(U)$ für alle $\lambda \in K$. Damit folgt die Behauptung. $\square$

7.3' Beispiele: Wir hatten bereits Beispiele aus der Elementargemometrie erwähnt. Eine sehr allgemeine Klasse von Beispielen linearer Abbildungen ergibt sich so:

Sei $\varphi : K^n \to K^m, (x_1, \dots, x_n) \mapsto (y_1, \dots, y_m)$ mit $y_i = \sum_{j=1}^{k} \alpha_{ij} x_j (i = 1, \dots, m)$, wobei $(\alpha_{ij})_{\substack{1 \leq i \leq m \\ 1 \leq j \leq n}}$ eine $(m \times n)$-Matrix ist. Dieses Beispiel ist, wie wir im nächsten Abschnitt sehen werden, sehr allgemein! Die Linearität von φ rechnet man sofort nach!

7.5 Bemerkung: Ganz entsprechend sieht man: Ist $\varphi : V \to W$ linear, $U \subset W$ ein Untervektorraum. Dann ist die Urbildmenge $\varphi^{-1}(U)$ ein Untervektorraum von V.

Beweis: Siehe Übungen!

7.6 Folgerung: Ist $\{e_1, \dots, e_n\}$ ein Erzeugendensystem des Vektorraums V, so ist $\{\varphi(e_1), \dots, \varphi(e_n)\}$ ein Erzeugendensystem des Bildvektorraums $\varphi(V)$.

Beweis: Übung.

7.7 Satz: Sei $\{e_1, \dots, e_n\}$ eine Basis des Vektorraums $V, \varphi : V \to W$ eine lineare Abbildung. Dann wird durch die Bildvektoren $w_1 = \varphi(e_1), \dots, w_n = \varphi(e_n)$ die Abbildung φ bereits eindeutig festgelegt.

Beweis: Sei $x \in V$ beliebig, $x = \sum_{i=1}^{n} \lambda_i e_i$ die eindeutige Darstellung des Vektors x durch die Basis $\{e_1, \dots, e_n\}$. Dann ist bei gegebenem φ

$$\begin{aligned} \varphi(x) &= \varphi\left(\sum_{i=1}^{n} \lambda_i e_i\right) \\ &= \sum_{i=1}^{n} \lambda_i \varphi(e_i) \end{aligned}$$

Sind umgekehrt nur die Vektoren $w_1, \dots, w_n \in W$ gegeben, so wird durch die obige Formel $\varphi(x)$ eindeutig definiert. Man rechnet leicht nach, dass die so definierte Abbildung linear ist. $\square$

Wir führen jetzt den folgenden wichtigen Begriff ein. Sei dafür wieder $\varphi : V \to W$ eine lineare Abbildung der Vektorräume V, W.

7.8 Definition und Satz: Der Kern der linearen Abbildung φ ist

$$\operatorname{Ker}(\varphi) := \{v \in V | \varphi(v) = 0\}.$$

$\operatorname{Ker}(\varphi)$ ist ein Untervektorraum des Vektorraums V.

Beweis: Seien $v_1, v_2 \in \operatorname{Ker}(\varphi), \lambda \in K$. Dann ist offenbar $\varphi(v_1 + v_2) = \varphi(v_1) + \varphi(v_2) = 0 + 0 = 0$, also auch $v_1 + v_2 \in \operatorname{Ker}(\varphi)$. Entsprechend ist $\varphi(\lambda v_1) = \lambda \varphi(v_1) = \lambda \cdot 0 = 0$ und daher auch $\lambda v_1 \in \operatorname{Ker}(\varphi)$. Endlich ist sicherlich der Nullvektor $0 \in \operatorname{Ker}(\varphi)$, da $\varphi(0) = 0$ ist. Es folgt daher, dass $\operatorname{Ker}(\varphi)$ ein Untervektorraum von V ist. $\square$

7.8' Bemerkung: Natürlich ist $\operatorname{Ker}(\varphi) = \varphi^{-1}(\{0\})$, insofern hätten wir uns in **7.8** auch auf **7.5** beziehen können.

7.9 Satz: (Dimensionsformel) Sei $\varphi : V \to W$ eine lineare Abbildung. Dann gilt: $\dim(V) = \dim \operatorname{Ker}(\varphi) + \dim \operatorname{Im}(\varphi)$, wobei insbesondere die eine Seite endlich ist genau wenn die andere Seite endlich ist.

Zuerst ein einfaches *Beispiel:* Sei $\varphi : K^2 \to K^2, (x_1, x_2) \mapsto (x_1, 0)$. Diese Abbildung (Projektion) ist offenbar linear. Man rechnet sofort nach: $\operatorname{Ker}(\varphi) = Ke_2, \operatorname{Im}(\varphi) = Ke_1$.

Beweis: Wir nehmen zunächst an, dass $\operatorname{Ker}(\varphi)$ und $\operatorname{Im}(\varphi)$ endlich dimensionale Vektorräume sind. Sei $\{e_1, \ldots, e_r\}$ eine Basis von $\operatorname{Ker}(\varphi)$. Weiter sei $\{\overline{e}_{r+1}, \ldots, \overline{e}_n\}$ eine Basis des Vektorraums $\varphi(V)$. Seien $e_{r+1}, \ldots, e_n$ Vektoren aus V mit $\varphi(ej) = \overline{e}_j \quad (j = r+1, \ldots, n)$.

Wir behaupten: $\{e_1, \ldots, e_n\}$ ist dann eine Basis von V: Sei nämlich $v \in V$ beliebig. Da $\varphi(v) \in \varphi(V)$ ist, existieren Skalare $\lambda_{r+1}, \ldots, \lambda_n$ so dass

$$\begin{aligned} \varphi(v) &= \sum_{i=r+1}^{n} \lambda_i \overline{e}_i \\ &= \sum_{i=r+1}^{n} \lambda_i \varphi(e_i) \end{aligned}$$

Daraus folgt mit Linearität von φ:

$$\varphi(v - \sum_{i=r+1}^{n} \lambda_i e_i) = 0$$

Also ist $v - \sum_{i=r+1}^{n} \lambda_i e_i \in \operatorname{Ker}(\varphi)$ nach Definition des Kerns. Daher existieren Skalare $\lambda_1, \ldots, \lambda_r \in K$ mit $\sum_{i=1}^{r} \lambda_i e_i = v - \sum_{i=r+1}^{n} \lambda_i e_i$. Daher folgt $\sum_{i=1}^{n} \lambda_i e_i = v$. Demnach ist $\{e_1, \ldots, e_n\}$ ein Erzeugendensystem von V.

Wir zeigen die lineare Unabhängigkeit des Systems $\{e_1, \ldots, e_n\}$ durch indirekten Beweis: Angenommen, das System $\{e_1, \ldots, e_n\}$ wäre nicht linear unabhängig. Dann bestünde eine nichttriviale Relation $\sum_{i=1}^{n} \lambda_i e_i = 0$. Daher folgt $\varphi\left(\sum_{i=1}^{n} \lambda_i e_i\right) = 0$. Andererseits ist

$$\begin{aligned} \varphi\left(\sum_{i=1}^{n} \lambda_i e_i\right) &= \sum_{i=1}^{n} \lambda_i \varphi(e_i) \\ &= \sum_{i=r+1}^{n} \lambda_i \overline{e}_i. \end{aligned}$$

Also folgt $\sum_{i=r+1}^{n} \lambda_i \overline{e}_i = 0$. Da nach Konstruktion des Systems $\{\overline{e}_{r+1}, \ldots, \overline{e}_n\}$ linear unabhängig ist, folgt, dass $\lambda_{r+1} = \cdots = \lambda_n = 0$ ist. Dann ergibt sich aber $\sum_{i=1}^{k} \lambda_i e_i = 0$.

Aber $\{e_1, \ldots, e_r\}$ war eine Basis des Vektorraums $\operatorname{Ker}(\varphi)$. Also folgt auch $\lambda_1 = \cdots = \lambda_r = 0$. Damit ist die lineare Unabhängigkeit gezeigt, $\{e_1, \ldots, e_n\}$ ist eine Basis von V. Wegen $r + (n-r) = n$ folgt $\dim \operatorname{Ker}(\varphi) + \dim \operatorname{Im}(\varphi) = \dim(V)$. Insbesondere ist $\dim(V)$ endlich, wenn $\dim \operatorname{Ker}(\varphi)$ und $\dim \operatorname{Im}(\varphi)$ endlich sind.

Sei jetzt umgekehrt vorausgesetzt, dass V ein endlichdimensionaler Vektorraum ist. Dann ist auch $\mathrm{Ker}(\varphi)$ als Teilraum endlichdimensional. Weiter ist $\varphi(V)$ endlichdimensional, da das Bild $\{\varphi(e_1), \dots, \varphi(e_n)\}$ einer Basis $\{e_1, \dots, e_n\}$ von V ein endliches Erzeugendensystem von $\varphi(V)$ ist. Dann ist aber die Vektorraum $\varphi(V)$ sicherlich endlichdimensional und wir sind genau in der am Anfang dieses Beweises betrachteten Situation. $\square$

8 Lineare Abbildungen II

Seien U, V, W Vektorräume über dem Körper K. Es seien $\varphi : V \to W$, $\psi : U \to V$ lineare Abbildungen.

8.1 Definition und Satz: Die zusammengesetzte Abbildung

$$\begin{array}{lcll} \varphi \circ \psi & : & U \to W, & \\ (\varphi \circ \psi)(x) & := & \varphi(\psi(x)) & \quad \text{für } x \in U \text{ beliebig} \end{array}$$

heißt Produkt oder Komposition von φ und ψ. Sie ist insbesondere wieder eine lineare Abbildung.

Beweis: Seien $x, y \in U$ Vektoren, $\lambda, \mu \in K$ Skalare. Dann gilt offenbar

$$\begin{array}{rcll} (\varphi \circ \psi)(\lambda x + \mu y) & = & \varphi(\psi(\lambda x + \mu y)) & \\ & = & \varphi(\lambda \psi(x) + \mu \psi(y)) & \quad \text{(wegen Linearität von } \psi) \\ & = & \lambda \varphi(\psi(x)) + \mu \varphi(\psi(y)) & \quad \text{(wegen Linearität von } \varphi) \\ & = & \lambda(\varphi \circ \psi)(x) + \mu(\varphi \circ \psi)(y) & \quad \text{(Def. der Komposition)} \end{array}$$

Daher ist $(\varphi \circ \psi)$ eine lineare Abbildung. □

8.2 Bemerkung: Man kann also lineare Abbildungen nur miteinander multiplizieren, wenn sie zueinander passen, d.h. wenn der Zielvektorraum der ersten Abbildung der Gegenstandsvektorraum der zweiten Abbildung ist.

Man kann übrigens lineare Abbildungen auch addieren, wenn sie zueinander passen (in anderem Sinn als oben!).

Seien nämlich

$$\varphi, \psi \quad : \quad V \to W$$

jetzt zwei lineare Abbildungen von V nach W.

8.3 Definition und Satz:

i) Die Summe

$$(\varphi + \psi)(x) : V \to W$$

ist definiert durch

$$(\varphi + \psi)(x) := \varphi(x) + \psi(x).$$

Sie ist wiederum eine lineare Abbildung.

ii) Entsprechend sei $\lambda \in K$ ein Skalar,

$$\lambda\varphi : V \to W$$

sei definiert durch

$$(\lambda\varphi)(x) := \lambda\varphi(x).$$

Auch $(\lambda\varphi)$ ist wieder eine lineare Abbildung.

Beweis: Dies kann als Übung dem Leser überlassen bleiben. □

8.4 Satz: Mit der obigen Addition und skalaren Multiplikation bildet die Menge $\mathrm{Hom}(V, W)$ aller linearen Abbildungen von V nach W selber einen K-Vektorraum.

Beweis: Wir hatten in **Kap. 3** die Menge $\mathrm{Abb}(V, W)$ *aller* (nicht nur der linearen) Abbildungen mit genau derselben Addition und skalaren Multiplikation versehen. Um zu sehen, dass $\mathrm{Hom}(V, W)$ ein K-Vektorraum ist, reicht es also aus zu zeigen, dass $\mathrm{Hom}(V, W)$ ein Untervektorraum von $\mathrm{Abb}(V, W)$ ist. Dazu reicht es, wie üblich, zu zeigen, dass $\mathrm{Hom}(V, W)$ in $\mathrm{Abb}(V, W)$ unter Summe und Multiplikation mit Skalaren abgeschlossen ist. Dies ist als Übung dem Leser überlassen. □

Eine besondere Rolle spielen die linearen Abbildungen (Homomorphismen) eines Vektorraums V in sich selber. Man spricht dann auch oft von *Endomorphismen* und schreibt statt $\mathrm{Hom}(V, V)$ oft auch $\mathrm{End}(V)$.

Offenbar kann man die Endomorphismen eines Vektorraums miteinander sowohl addieren als auch multiplizieren. Es gelten die üblichen Rechenregeln, dabei seien $\varphi, \psi, \chi, \ldots \in \mathrm{End}(V)$.

$$\begin{array}{rcl} (\varphi \circ \psi) \circ \chi &=& \varphi \circ (\psi \circ \chi) \\ \varphi \circ (\psi + \chi) &=& \varphi \circ \psi + \varphi \circ \chi \\ (\varphi + \psi) \circ \chi &=& \varphi \circ \chi + \psi \circ \chi \end{array}$$

Weiter ist $\varphi \circ 0 = 0 \circ \varphi = 0$ für die Nullabbildung 0.

Endlich gibt es ein neutrales Element der Multiplikation, die identische Abbildung

$$\begin{array}{rcccl} \mathrm{id} &:& V &\to& V \\ && x &\mapsto& x \end{array}$$

Es gilt

$$\varphi \circ \mathrm{id} = \mathrm{id} \circ \varphi = \varphi$$

für alle $\varphi \in \mathrm{End}(V)$.

Allerdings ist die Multiplikation von Endomorphismen i. A. nicht kommutativ. D. h. es gilt *nicht immer* $\varphi \circ \psi = \psi \circ \varphi$.

8.5 Satz: Die Menge der Endomorphismen eines Vektorraums bildet mit obiger Addition und Multiplikation einen Ring.

Besonders wichtig sind die bijektiven linearen Abbildungen $\varphi : V \to W$. Wir können die Umkehrabbildung $\varphi^{-1} : W \to V$ betrachten, so dass also gilt:
Es ist für $w \in W$, $v \in V$ $\varphi^{-1}(w) = v$ genau dann, wenn $\varphi(v) = w$ ist.

8.6 Satz: Ist

$$\varphi : V \to W$$

eine bijektive lineare Abbildung, dann ist die inverse Abbildung

$$\varphi^{-1} : W \to V$$

ebenfalls eine lineare Abbildung. Es gilt

$$\begin{aligned} \varphi \circ \varphi^{-1} &= \mathrm{id}_W \\ \varphi^{-1} \circ \varphi &= \mathrm{id}_V . \end{aligned}$$

Beweis: Wir zeigen nur die Linearität der Umkehrabbildung φ^{-1}. Seien $w_1, w_2 \in W$. Da φ bijektiv ist, existieren eindeutige Urbilder $v_1, v_2 \in V$, $\varphi(v_i) = w_i \quad i = 1, 2$.
Dann ist aber $\varphi(v_1 + v_2) = w_1 + w_2$, also

$$\begin{aligned} \varphi^{-1}(w_1 + w_2) &= v_1 + v_2 \\ &= \varphi^{-1}(w_1) + \varphi^{-1}(w_2). \end{aligned}$$

Entsprechend zeigt man

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

□

8.7 Definition und Satz: Die Menge der bijektiven Endomorphismen eines Vektorraums ist die sog. allgemeine lineare Gruppe $\mathrm{GL}(V)$. Sie ist abgeschlossen gegenüber Produkten und Inversenbildung.

8.8 Bemerkung: Die Bezeichnung $\mathrm{GL}(V)$ kommt aus dem Englischen, *general linear group* wörtlich genau entsprechend.

Beweis von Satz 8.7: Mit φ, ψ bijektiv und linear, ist offenbar auch $\varphi \circ \psi$ bijektiv und linear, gehört also ebenfalls zu $\mathrm{GL}(V)$.
Ist φ bijektiv linear, so auch φ^{-1} (Beweis dafür Übung). Man überzeugt sich leicht, dass $(\varphi^{-1})^{-1} = \varphi$, falls $\varphi \in \mathrm{GL}(V)$. Endlich ist die identische Abbildung $\mathrm{id} \in \mathrm{GL}(V)$.□

8.9 Bemerkung: Die Menge $(\mathrm{GL}(V), \cdot)$, zusammen mit der Multiplikation linearer Abbildungen bildet eine sog. *Gruppe*.
Gruppen sind fundamentale algebraische Strukturen, die aber auch an vielen Stellen in der Physik eine große Rolle spielen, etwa als Symmetriegruppen oder z. B. die sog. Lorentz-Gruppe in der Relativitätstheorie. In **Kap. 13** wird mehr über Gruppen verraten werden.

9 Lineare Abbildungen und Matrizen

Will man Punkte im Raum festlegen, so braucht man ein Koordinatensystem, entsprechend legt man in einem abstrakten K-Vektorraum V die Vektoren durch ihre Komponenten bezüglich einer fixierten Basis $\{e_1, \dots, e_n\}$ fest.

Will man jetzt eine lineare Abbildung

$$\varphi : V \to W$$

zwischen Vektorräumen V, W rechnerisch konkret beschreiben, so geht man so vor:

Wir fixieren Basen $\{e_1, \dots, e_n\}$ von V, $\{f_1, \dots, f_m\}$ von W. Wir betrachten die Bildvektoren $\varphi(e_j)$ für $j = 1, \dots, n$. Da $\varphi(e_j) \in W$ und $\{f_1, \dots, f_m\}$ eine Basis von W ist, können wir schreiben:

$$\varphi(e_j) = \sum_{i=1}^{m} a_{ij} f_i \quad (j = 1, \dots, n)$$

Es ergibt sich also ein „rechteckiges“ Schema von Skalaren (zum Beispiel reellen Zahlen, wenn $K = \mathbb{R}$ ist):

$$\begin{pmatrix} a_{11} & \dots & a_{1n} \\ a_{21} & \dots & a_{2n} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix}$$

oder kürzer:

$$(a_{ij})_{\substack{1 \le i \le m \\ 1 \le j \le n}}$$

Ein derartiges rechteckiges System nennen wir eine *Matrix* (mit „Eingängen“ oder auch Komponenten aus dem Körper K). Wir hatten derartige rechteckige Systeme bereits in anderem Zusammenhang in **Kap. 6** kennengelernt.

Wir schreiben

$$(a_{ij})_{\substack{1 \le i \le m \\ 1 \le j \le n}} =: \mathrm{Matrix}(\varphi; \{e_j\}_{j=1,\dots,n}; \{f_i\}_{i=1,\dots,m})$$

In Worten: $A = (a_{ij})$ ist die Matrix der linearen Abbildung $\varphi : V \to W$ bezüglich der Basen $\{e_1, \dots, e_n\}$ von V und $\{f_1, \dots, f_m\}$ von W.

Es ist plausibel, dass die Matrix A nicht nur von der linearen Abbildung φ, sondern auch von der Wahl der Basen abhängen wird. Wie, das werden wir weiter unten ausrechnen!

Die Vektoren des K^n, $(a_{i1}, \dots, a_{in})$, heißen *Zeilen* der Matrix A,
die Vektoren $(a_{1j}, a_{2j}, \dots, a_{mj}) \in K^m$ heißen die *Spalten* der Matrix A.

Die Wahl der Indizes oben ist so eingerichtet, dass folgendes gilt:

9.1 Festlegung: Die Komponenten des Bildes des j-ten Basisvektors e_j findet man in der j-ten Spalte der Matrix A.

Wie gesagt handelt es sich hierbei nur um eine Festlegung, aber sie ist sehr wichtig und man sollte sie sich ein für allemal merken.

Sei jetzt $x = \sum_{j=1}^{n} \xi_j e_j$ ein beliebiger Vektor aus V. Wir berechnen

$$\begin{aligned}
\varphi(x) &= \sum_{j=1}^{n} \xi_j \varphi(e_j) \quad \text{(mit Linearität von } \varphi\text{)} \\
&= \sum_{j=1}^{n} \xi_j \left(\sum_{i=1}^{m} a_{ij} f_i \right) \quad \text{(nach Definition)} \\
&= \sum_{i=1}^{m} \left(\sum_{j=1}^{n} a_{ij} \xi_j \right) f_i \\
&= \sum_{i=1}^{m} \eta_i f_i \quad \text{mit } \eta_i = \sum_{j=1}^{n} a_{ij} \xi_j \quad (i = 1, \dots, m)
\end{aligned}$$

Damit sehen wir noch einmal, was aber bereits vorher klar war, dass die Festlegung der Bildvektoren $\varphi(e_1), \dots, \varphi(e_n)$ bereits die lineare Abbildung φ und natürlich auch die Matrix $M(\varphi; \dots)$ festlegt.

9.2 Definition und Satz: Die Menge der $m \times n$-Matrizen $A = (a_{ij})_{\substack{1 \le i \le m \\ 1 \le j \le n}}$ mit Komponenten aus K, die wir mit $\mathbb{M}(m \times n; K)$ bezeichnen, bildet einen K-Vektorraum der Dimension $(m \cdot n)$.

Beweis: Natürlich werden wir dabei zwei $m \times n$-Matrizen $A = (a_{ij})$, $B = (b_{ij})$ mit $1 \le i \le m$, $1 \le j \le n$ so addieren:

$$A + B := (c_{ij})_{\substack{1 \le i \le m \\ 1 \le j \le n}} \quad \text{mit } c_{ij} = a_{ij} + b_{ij}$$

Entsprechend multipliziert man mit Skalaren so:

$$\lambda (a_{ij})_{\substack{1 \le i \le m \\ 1 \le j \le n}} = (\lambda a_{ij})_{\substack{1 \le i \le m \\ 1 \le j \le n}}$$

Damit wird $\mathbb{M}(m \times n; K)$ selber zu einem Vektorraum. Eine Basis wird zum Beispiel von dem System der Matrizen E_{ij} $(1 \le i \le m;\ 1 \le j \le n)$ gebildet, wobei E_{ij} diejenige Matrix ist, die lediglich an der (i,j)-ten Stelle eine 1 stehen hat, sonst lauter Nullen, also:

$$(E_{ij})_{\nu,\mu} = \delta_{i\nu} \delta_{j\mu}$$

mit Hilfe des sogenannten Kroneckersymbols ($\delta_{\alpha\beta} = 1$ für $\alpha = \beta$, 0 sonst). Die Anzahl der Basisvektoren $\{E_{ij}\}$ ist gerade $(m \cdot n)$. Damit folgt die Behauptung. $\square$

Sei, wie oben, $\{e_1, \ldots, e_n\}$ eine feste Basis des Vektorraums V, $\{f_1, \ldots, f_m\}$ eine Basis von W. Unsere obigen Ergebnisse können wir jetzt so zusammenfassen:

9.3 Satz: Die obige Zuordnung

$$\varphi \longmapsto \mathrm{Matrix}(\varphi; \{e_j\}; \{f_i\})$$

definiert einen Isomorphismus von K-Vektorräumen

$$\mathrm{Hom}_K(V, W) \longrightarrow \mathbb{M}(m \times n; K) \quad .$$

Da wir lineare Abbildungen zusätzlich, falls sie zueinander passen, auch multiplizieren können, wird man sich fragen, ob das auch für Matrizen richtig ist.

Sei etwa $\varphi : V \to W$ wie oben, $\psi : U \to V$ eine weitere lineare Abbildung zwischen den Vektorräumen U und V.

$\{e_1, \ldots, e_n\}$ und $\{f_1, \ldots, f_m\}$ seien wie gehabt Basen von V und W.
Weiter sei $\{d_1, \ldots, d_r\}$ eine Basis von U. Es sei

$$\psi(d_k) = \sum_{j=1}^{n} b_{jk} e_j \quad (k = 1, \ldots, r)$$

und also

$$\mathrm{Matrix}(\psi; \{d_k\}; \{e_j\}) = (b_{jk}) =: B$$

Insgesamt ergibt sich

$$\begin{aligned}
(\varphi \circ \psi)(d_k) &= \varphi(\psi(d_k)) \\
&= \varphi(\sum_{j=1}^{n} b_{jk} e_j) \\
&= \sum_{j=1}^{n} b_{jk} \varphi(e_j) \quad \text{(Linearität von } \varphi) \\
&= \sum_{j=1}^{n} b_{jk} \left(\sum_{i=1}^{m} a_{ij} f_i \right) \\
&= \sum_{i=1}^{m} \left(\sum_{j=1}^{n} a_{ij} b_{jk} \right) f_i \\
&= \sum_{i=1}^{m} c_{ik} f_i \quad \text{mit } (c_{ik})_{\substack{1 \le i \le m \\ 1 \le k \le r}} = \mathrm{Matrix}(\varphi \circ \psi; \{d_k\}; \{f_i\})
\end{aligned}$$

Es liegt jetzt nahe, die sich ergebende Matrix $(c_{ik}) =: C$ als Produkt der Matrizen A, B zu definieren.

9.4 Definition: Seien $A = (a_{ij})_{\substack{1\le i\le m\\1\le j\le n}} \in \mathbb{M}(m \times n; K)$, $B = (b_{jk})_{\substack{1\le j\le n\\1\le k\le r}} \in \mathbb{M}(n \times r; K)$. Dann ist $C = (c_{ik})_{\substack{1\le i\le m\\1\le k\le r}} \in \mathbb{M}(m \times r; K)$ mit $c_{ik} = \sum_{j=1}^{n} a_{ij}b_{jk} \quad (1 \le i \le m; 1 \le k \le r)$ das Produkt $A \cdot B$ von A und B.

9.5 Bemerkung: Um also die (i,k)-te Komponente im Produkt zu erhalten, multipliziere man komponentenweise die i-te Zeile von A, $(a_{i1}, \dots, a_{in})$, mit der k-ten Spalte von B, $(b_{1k}, b_{2k}, \dots, b_{nk})$. Man beachte, man erhält jeweils Vektoren mit n Komponenten, die man komponentenweise multiplizieren kann.

Beispiel: Es sei gleich darauf hingewiesen: Die Matrixmultiplikation ist *nicht* kommutativ: $A = \left(\begin{smallmatrix}1&1\\0&1\end{smallmatrix}\right)$, $B = \left(\begin{smallmatrix}1&0\\1&1\end{smallmatrix}\right)$.

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

Für Addition und Multiplikation von Matrizen gelten folgenden Rechenregeln, dabei ist vorausgesetzt, dass die unten auftretenden Ausdrücke *alle* definiert sind, d. h. die Matrizen die richtigen Formate haben.

$$\begin{aligned} A+B &= B+A \\ (A+B)+C &= A+(B+C) \\ A(B+C) &= AB+AC \\ (A+B)C &= AC+BC \\ (AB)C &= A(BC) \end{aligned}$$

Betrachtet man speziell den Vektorraum $\mathbb{M}(n \times n; K)$ der $(n \times n)$-Matrizen (sog. quadratische Matrizen, für die wir auch einfacher $\mathbb{M}(n; K)$ schreiben), so können Addition und Multiplikation uneingeschränkt ausgeführt werden.

$\mathbb{M}(n; K)$ wird dann wiederum ein Ring, ganz entsprechend dem Ring der Endomorphismen des Vektorraums V.

9.6 Satz: $\{e_1, \dots, e_n\}$ sei eine beliebige feste Basis des Vektorraumes V. Es sei

$$\begin{aligned} \mathrm{End}_K(V) &\longrightarrow \mathbb{M}(n; K) \\ \varphi &\longmapsto \mathrm{Matrix}(\varphi; \{e_i\}; \{e_i\}) \end{aligned}$$

die Abbildung, die jedem Endomorphismus φ die Matrix bezüglich der Basis $\{e_i\}$ (beidemale verwendet) zuordnet. Dies ist ein K-Isomorphismus von Vektorräumen, der zusätzlich die Multiplikation respektiert, also, wie man sagt, sogar ein Isomorphismus der Ringstruktur ist.

Beweis: Es ist bereits alles gezeigt. Insbesondere folgt noch einmal in dieser Situation, dass die Matrixmultiplikation auch assoziativ ist. $\square$

Sei noch einmal $\varphi : V \to W$, linear, $A = (a_{ij}) = \mathrm{Matrix}(\varphi; \{e_j\}; \{f_i\})$ wie oben. Wir berechnen $\varphi(V) = \mathrm{Im}(\varphi)$, das Bild von φ. Wir wissen bereits, dass dann $\{\varphi(e_1), \dots, \varphi(e_n)\}$ ein Erzeugendensystem von $\varphi(V)$ ist. Daher ist ein maximales Teilsystem linear unabhängiger Vektoren eine Basis von $\varphi(V)$.

9.7 Definition: Der *Zeilenrang* einer $(m \times n)$-Matrix $(\alpha_{ij})_{\substack{1\le i\le m\\ 1\le j\le n}}$ ist die Dimension des von den Zeilen $\{(\alpha_{i1}, \dots, \alpha_{in}) | 1 \le i \le m\}$ aufgespannten Untervektorraumes vom K^n, oder äquivalent: die Anzahl eines maximalen linear unabhängigen Teilsystems der Zeilen.

Wörtlich entsprechend definiert man den *Spaltenrang* der Matrix (α_{ij}).

Wir haben uns also oben überlegt:

9.8 Satz: Die Dimension des Bildes $\varphi(V)$ einer linearen Abbildung $\varphi : V \to W$ ist gerade der (Spalten)rang einer beliebigen φ zugeordneten Matrix (α_{ij}).

9.9 Bemerkung: Wir werden weiter unten einsehen, dass für eine beliebige Matrix A gilt: $\mathrm{Spaltenrang}(A) = \mathrm{Zeilenrang}(A)$. Daher schreiben wir einfach $\mathrm{rang}(A)$ für diese Zahl.

Ehe wir die Abhängigkeit einer Matrix von den gewählten Basen berechnen, lösen wir die folgende einfache Aufgabe.

Sei V ein Vektorraum, $\{e_1, \dots, e_n\}$ und $\{e'_1, \dots, e'_n\}$ seien Basen, wir betrachten die identische Abbildung

$$\mathrm{id} : V \longrightarrow V$$

und wollen

$$\mathrm{Matrix}(\mathrm{id}; \{e_1, \dots, e_n\}; \{e'_1, \dots, e'_n\})$$

berechnen.

Es sei $e_j = \sum_{i=1}^n \gamma_{ij} e'_i$ $(j = 1, \dots, n)$ eine Darstellung der Basisvektoren e_j durch die e'_i.

Offenbar ist

$$C = (\gamma_{ij}) = \mathrm{Matrix}(\mathrm{id}; \{e_j\}; \{e'_i\})$$

Ganz entsprechend zeigt man: Mit $e'_j = \sum_{i=1}^n \delta_{ij} e_i$ und $D := (\delta_{ij})_{1\le i,j\le n}$ ist

$$\mathrm{Matrix}(\mathrm{id}; \{e'_j\}; \{e_i\}) = D$$

Dann ist $C \cdot D$ die Matrix der zusammengesetzten linearen Abbildung $\mathrm{id} : V \to V$ bezüglich der Basis $\{e'_1, \dots, e'_n\}$, also $C \cdot D = \mathrm{Matrix}(\mathrm{id} : V \to V; \{e'_j\}; \{e'_i\})$. Daher ist $C \cdot D = E$, die Einheitsmatrix. Ganz entsprechend zeigt man $D \cdot C = E$. Das heißt, es ist $D = C^{-1}$.

Jetzt betrachten wir allgemeiner folgende Situation: $\varphi : V \to W$ sei eine lineare Abbildung, $\{e_1, \dots, e_n\}$ und $\{e'_1, \dots, e'_n\}$ Basen von V, $\{f_1, \dots, f_m\}$ und $\{f'_1, \dots, f'_m\}$ Basen von W.

Wir betrachten

$$(V;\{e_i'\}) \xrightarrow{\text{id}} (V;\{e_i\}) \xrightarrow{\varphi} (W;\{f_i\}) \xrightarrow{\text{id}} (W;\{f_i'\})$$

Bezeichnet noch

$$B = \text{Matrix}(\text{id};\{f_i\};\{f_i'\}) \quad ,$$

d. h. explizit gilt

$$f_j = \sum_{i=1}^{m} \beta_{ij} f_i' \quad ,$$

so folgt

$$M(\varphi;\{e_i'\};\{f_i'\}) = M(\text{id};\{f_i\};\{f_i'\}) \cdot M(\varphi;\{e_i\};\{f_i\}) \cdot M(\text{id};\{e_i'\};\{e_i\})$$

d. h. (Transformationsgesetz für Matrizen)

$$A' = (BAC^{-1}) \quad \text{mit } A' = M(\varphi;\{e_i'\};\{f_i'\})$$

Ein besonders interessanter Spezialfall ist der Fall eines *Endomorphismus* $\varphi : V \to V$ (d. h. $W = V$). Hier möchte man oft die Basis im Ziel- und Gegenstandsraum gleich wählen, d. h. $\{e_i\} = \{f_i\}$ und entsprechend $\{e_i'\} = \{f_i'\}$.

Obiges *Transformationsgesetz* für Basisänderungen sieht dann so aus (Transformationsformel für Matrizen von Endomorphismen):

$$A' = BAB^{-1}$$

Dabei ist also B gegeben durch

$$e_j = \sum_{i=1}^{n} \beta_{ij} e_i' \quad (j = 1,\ldots,n)$$
$$B = (\beta_{ij})_{1\le i,j\le n}$$

Wir fassen die gewonnenen Ergebnisse zusammen:

9.10 Satz: *(Transformationsverhalten von Matrizen bei Basisänderung)*

i) Ist $\varphi : V \to W$ eine lineare Abbildung, $\{e_1,\ldots,e_n\}$ und $\{e_1',\ldots,e_n'\}$ Basen von V, entsprechend $\{f_1,\ldots,f_m\}$ und $\{f_1',\ldots,f_m'\}$ Basen von W, $A := \text{Matrix}(\varphi;\{e_j\};\{f_i\})$ und $A' := \text{Matrix}(\varphi;\{e_j'\};\{f_i'\})$. Weiter seien die Übergangsmatrizen B und C gegeben durch $B = \text{Matrix}(\text{id};\{f_i\};\{f_i'\})$ und $C = \text{Matrix}(\text{id};\{e_j\};\{e_j'\})$, d. h. es gelten die Beziehungen

$$e_j = \sum_{i=1}^{n} \gamma_{ij} e_i' \quad (j = 1,\ldots,n) \qquad \text{und} \qquad C = (\gamma_{ij}) \quad ,$$
$$f_j = \sum_{i=1}^{n} \beta_{ij} f_i' \quad (j = 1,\ldots,n) \qquad \text{und} \qquad B = (\beta_{ij})$$

Dann ist $A' = BAC^{-1}$.

ii) Ist speziell $V = W$, $\varphi : V \to V$ ein Endomorphismus, so spezialisiert sich die obige Formel zu $A' = BAB^{-1}$.

Das Normalformenproblem

Man kann sich fragen, wie man durch geschickte Wahl von Basen einen Vektorraumhomomorphismus bzw. -endomorphismus auf möglichst einfache Gestalt bringen kann?

Sei zunächst $\varphi : V \to W$ ein beliebiger Homomorphismus. Wir wählen wie folgt eine Basis von V, $\dim(V) = n$: $\{e_1, \ldots, e_n\}$, dabei sei $\{e_{r+1}, \ldots, e_n\}$ eine Basis von $\operatorname{Ker}(\varphi)$. Daher ist $\{\varphi(e_1), \ldots, \varphi(e_r)\}$ eine Basis des Untervektorraumes $\operatorname{Im}(\varphi) \subset W$, die wir beliebig zu einer Basis von W ergänzen. Also $f_j = \varphi(e_j)$ $(j = 1, \ldots, r)$. Man sieht sofort:

$$M(\varphi; \{e_i\}; \{f_i\}) = \left(\begin{array}{c|c} E_r & 0 \\ \hline 0 & 0 \end{array}\right)$$

dabei haben wir die $(m \times n)$-Matrix oben in offensichtlicher Weise in Blöcke unterteilt, insbesondere steht E_r für die $(r \times r)$-Einheitsmatrix.

Man sieht also, das Ergebnis hängt einzig und allein vom Rang der Abbildung, d. h. von $\dim(\operatorname{Im}(\varphi))$, ab. Speziell, ist $\varphi : V \to W$ ein Isomorphismus, so ist bei Wahl der Basen $\{e_1, \ldots, e_n\}$ (hier ganz beliebig und $\{f_1, \ldots, f_n\}$ mit $\varphi(e_j) = f_j$)

$$M(\varphi; \{e_1, \ldots, e_n\}; \{f_1, \ldots, f_n\}) = E_n$$

9.12 Satz: Ist $\varphi : V \to W$ eine lineare Abbildung vom Rang r, so ist bei geeigneter Wahl von Basen $\{e_j\}$ von V, $\{f_i\}$ von W:

$$M(\varphi; \{e_i\}; \{f_i\}) = \begin{pmatrix} E_r & 0 \\ 0 & 0 \end{pmatrix}$$

in Blockschreibweise.

In gewissem Sinne besitzt eine solche Normalform nicht viel Aussagekraft. Ganz anders wird es mit dem *Normalformenproblem für Endomorphismen* stehen. Dies werden wir im Kapitel über Eigenwerte und Eigenvektoren etwas näher kennenlernen und es wird auf die sogenannte Jordansche Normalform führen, die eines der starken Hilfsmittel der linearen Algebra darstellt.

Ergänzung: $\varphi : V \to W$, $\{e_j\}$, $\{f_i\}$ Basen; $M(\varphi; \{e_j\}; \{f_i\}) = (\alpha_{ij})_{\substack{1\le i\le m\\ 1\le j\le n}}$

$$x \in V\ ,\quad x = \sum_{j=1}^{n} \xi_j e_j \Longrightarrow$$

$$\varphi(x) = \sum_{i=1}^{m} \left(\sum_{j=1}^{n} \alpha_{ij}\xi_j \right) f_i$$

$$\eta_i := \sum_{j=1}^{n} \alpha_{ij}\xi_j \quad (i = 1, \dots, m)$$

$$\begin{pmatrix} \eta_1 \\ \vdots \\ \eta_m \end{pmatrix} = (\alpha_{ij}) \begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix}$$

in Matrixschreibweise. Man kann die lineare Abbildung also ohne weiteres als Matrixmultiplikation schreiben, wenn man die Komponenten der Vektoren als Spalten schreibt.

Wir kommen zurück zum Begriff des Ranges einer Matrix A.

9.12a Satz: Spaltenrang und Zeilenrang einer Matrix A ändern sich nicht bei Änderung der Basen. Es gilt

$$\text{Spaltenrang}(A) = \text{Zeilenrang}(A)$$

Beweis: Eine Basisänderung der Matrix A bedeutet den Übergang von A zur Matrix $A' = BAC$ mit invertierbaren Matrizen B, C. Nun stellen A bzw. $A' = BAC$ dieselbe lineare Abbildung φ bezüglich verschiedener Basen des K^n bzw. des K^m dar. Es ist daher

$$\begin{aligned} \text{Spaltenrang}(A) &= \dim \varphi(K^n) \\ \text{Spaltenrang}(BAC) &= \dim \varphi(K^n) \end{aligned}$$

Daher ist aber auch

$$\text{Spaltenrang}(A) = \text{Spaltenrang}(BAC) \quad .$$

Damit folgt die Aussage über den Spaltenrang.

Wegen $\text{Zeilenrang}(A) = \text{Spaltenrang}(A^t)$ ergibt sich dann leicht die entsprechende Aussage für den Zeilenrang.

Durch geschickte Wahl von B, C können wir endlich erreichen, dass $A' = BAC$ Normalform hat, d. h.

$$A' = \begin{pmatrix} E_r & 0 \\ 0 & 0 \end{pmatrix}$$

wie in **9.11** ist.

Da aber sofort klar ist, dass gilt

$$\text{Spaltenrang}(A') = \text{Zeilenrang}(A') = r \quad ,$$

folgt nach dem schon gezeigten auch

$$\text{Spaltenrang}(A) = \text{Zeilenrang}(A) \quad .$$

□

9.13 Proposition: Elementare Zeilen- bzw. Spaltenumformungen ändern den Rang einer Matrix nicht.

Beweis: Wir hatten in **6.5** gesehen: Elementare Zeilenumformungen ändern den Zeilenrang nicht. Durch Transposition gilt gleiches für den Spaltenrang. □

9.14 Bemerkungen:

i) Man könnte auch direkt „zu Fuß“ nachrechnen, dass Zeilenumformungen *nicht* den Spaltenrang ändern. Da man mittels Zeilenumformungen eine Matrix auf Stufenform A' im Sinn von **6.5** bringen kann, genügt es noch zu beweisen, dass

$$\text{Spaltenrang}(A') = \text{Zeilenrang}(A')$$

gilt, was unmittelbar klar ist.

ii) Es ist oft nützlich zu wissen, dass sich elementare Zeilen- bzw. Spaltenumformungen durch Links- bzw. Rechtsmultiplikation mit sogenannten Elementarmatrizen ergeben. Sei für $1 \leq i, j \leq n, i \neq j$ etwa $C_{ij}(\lambda)$ folgende $(n \times n)$-Matrix:

$$C_{ij}(\lambda) := (E_n + \lambda E_{ij})$$

(d. h. $C_{ij}(\lambda)$ hat Einsen auf der Diagonalen, λ an der Stelle (i, j), sonst Nullen.

Bezeichnet a_k den k-ten Spaltenvektor von A, a'_k den k-ten Spaltenvektor von $AC_{ij}(\lambda)$ $(1 \leq k \leq n)$, so ergibt sich:

$$\begin{aligned} a'_k &= a_k \quad \text{für alle } k \neq j \\ a'_j &= a_j + \lambda a_i \end{aligned}$$

D. h. A' entsteht aus A durch Addition des λ-fachen der i-ten Spalte zur j-ten Spalte.

Übung: Man zeige die entsprechende Aussage für elementare Zeilenumformungen.

10 Linearformen, der duale Vektorraum

Sei V ein K-Vektorraum, wir betrachten lineare Abbildungen $f : V \to K$ von V in den 1-dimensionalen Vektorraum K. Diese sogenannten *Linearformen* auf V bilden selber einen Vektorraum $\mathrm{Hom}(V, K)$ (siehe **Kap. 8**), den Dualraum von V, für den wir auch kurz V^* schreiben.

10.1 Beispiel: Sei $V = K^n$, eine Linearform wird dann durch einen Ausdruck

$$f(x_1, \ldots, x_n) = \sum_{i=1}^{n} a_i x_i, \quad \text{mit } a_1, \ldots, a_n \in K \quad ,$$

gegeben.

10.2 Proposition: Sei $f : V \to K$ eine Linearform $\neq 0$. Dann ist $\mathrm{Ker}(f)$ eine Hyperebene in V, das heißt, ein Untervektorraum der Dimension $n - 1$.

Beweis: Ist $f \neq 0$, dann ist $f(V) \neq 0$, also $f(V) = K$, d.h. $\dim \mathrm{Im}(f) = 1$. Wegen der Dimensionsformel gilt

$$\dim \mathrm{Ker}(f) + \dim \mathrm{Im}(f) = n \quad ,$$

also $\dim \mathrm{Ker}(f) = (n - 1)$. Das war zu zeigen. □

10.3 Bemerkung:

i) Ist die Linearform also wie oben von der Form

$$f(x_1, \ldots, x_n) = \sum_{i=1}^{n} a_i x_i \quad ,$$

so ist $\mathrm{Ker}(f)$ gerade die durch

$$\sum_{i=1}^{n} a_i x_i = 0$$

gegebene Hyperebene im K^n.

ii) Oft schreibt man auch statt $f(v)$ für eine Linearform für $f \in V^*$, $v \in V$ $\langle v, f \rangle$. Dabei möchte man unterstreichen, dass man die Vektoren aus V und V^* wie ein Skalarprodukt aus Vektoren miteinander paart.

10.4 Definition und Satz: Sei V ein endlich-dimensionaler K-Vektorraum mit einer Basis $\{e_1, \dots, e_n\}$. Durch $f_i(e_j) = \delta_{ij}$ $(1 \leq i, j \leq n)$ werden dann Linearformen $f_1, \dots, f_n$ von V^* gegeben. $\{f_1, \dots, f_n\}$ ist die zu $\{e_1, \dots, e_n\}$ sogenannte *duale Basis von* V^*.

Beweis: Zu zeigen ist die bei dieser Vorgabe der $f_1, \dots, f_n$, die offenbar wohldefiniert sind, lineare Unabhängigkeit und Erzeugendeneigenschaft.

Wäre $\sum_{i=1}^n a_i f_i = 0$ in V^*, so folgte sofort

$$\left(\sum_{i=1}^n a_i f_i\right)(e_j) = a_j = 0 \quad \text{für } j = 1, \dots, n \quad ,$$

was gerade für die lineare Unabhängigkeit der $\{f_1, \dots, f_n\}$ in V^* zu zeigen ist.

Sei φ eine beliebige Linearform auf V. Dann gilt (Übung):

$$\varphi = \sum_{i=1}^n \varphi(e_i) f_i \quad .$$

Daher ist $\{f_1, \dots, f_n\}$ eine Basis von V^* □

10.5 Folgerung: Ist V endlich-dimensional, so ist $\dim(V) = \dim(V^*)$.

10.6 Definition: Sei V endlich-dimensionaler Vektorraum, V^* der Dualraum. Sei $W \subset V$ ein Teilraum. Dann bezeichnet

$$W^\perp := \{f \in V^* | f(w) = 0 \text{ für alle } w \in W\} \quad .$$

Entsprechend sei für einen Teilraum U von V^*:

$$U^\perp := \{v \in V | f(v) = 0 \text{ für alle } f \in U\} \quad .$$

10.7 Proposition: Es sind $W^\perp$ und $U^\perp$ Untervektorräume von V^* bzw. von V. Es gilt

i) $\dim(W)^\perp = \dim(V) - \dim(W)$

ii) $\dim(U)^\perp = \dim(V) - \dim(U)$

iii) $(W^\perp)^\perp = W, \quad (U^\perp)^\perp = U$

Beweis: Dass $U^\perp$, $W^\perp$ Untervektorräume sind, ist eine leichte Übung.

Zu (i) Sei $\{e_1, \dots, e_r\}$ eine Basis von W, die wir zu einer Basis $\{e_1, \dots, e_n\}$ von V ergänzen. Sei $\{f_1, \dots, f_n\}$ die dazu duale Basis von V^*, sodass also gilt: $f_i(e_j) = \delta_{ij}$ $(1 \leq i, j \leq n)$.

Dann ist (Übung) $\{f_{r+1}, \dots, f_n\}$ eine Basis von $(W)^\perp$. Damit folgt (i). (ii) und (iii) werden unten nachgewiesen. Zuerst zeigen wir dafür den

10.8 Satz: Sei V ein K-Vektorraum, V^* sein Dualraum und $(V^*)^*$ der sogenannte Bidualraum. Dann besteht eine kanonische injektive lineare Abbildung

$$\begin{cases} V & \to & (V^*)^* \\ v & \mapsto & (\varphi : V^* \to K) \text{ mit } \varphi(f) := f(v) \text{ für } f \in V^* \end{cases} \quad .$$

Ist V endlich-dimensional, so ist die kanonische Abbildung $V \to (V^*)^*$ eine Isomorphie.

Beweis: Zunächst ist die Zuordnung $v \mapsto \varphi$ linear, denn $(v_1 + v_2) \mapsto \varphi$ mit

$$\varphi(f) = f(v_1 + v_2) = f(v_1) + f(v_2) = \varphi_1(f) + \varphi_2(f) \quad ,$$

wobei $v_1 \mapsto \varphi_1$, $v_2 \mapsto \varphi_2$.

Weiter ist $V \to (V^*)^*$ eine injektive Abbildung. Angenommen, $v \mapsto \varphi = 0$ in $(V^*)^*$. Dann ist also $\varphi(f) = f(v) = 0$ für alle $f \in V^*$. Das geht nur, wenn $v = 0$ (sonst findet man sofort eine Linearform f mit $f(v) \neq 0$). Daher ist $V \to (V^*)^*$ injektiv.

Ist V zusätzlich endlich-dimensional, so ist $\dim(V) = \dim(V^*) = \dim((V^*)^*)$ nach **9.2**.

Dann ist aber die injektive Abbildung $V \to (V^*)^*$ wegen Gleichheit der Dimensionen sogar ein Isomorphismus. □

Jetzt können wir auch den Beweis von Satz **10.7** beenden.

Zu (ii) Wir betrachten $U \subset V^*$ und $(U)^\perp$ in $(V^*)^*$ (zunächst nicht in V !)

$$(U)^\perp = \{\varphi \in (V^*)^* | \varphi(f) = 0 \quad \text{für alle } f \in U\} \quad .$$

Nach dem bereits gezeigten Teil (i) folgt:

$$\dim(U)^\perp = \dim(V^*) - \dim(U) = \dim(V) - \dim(U) \quad .$$

Da die kanonische Abbildung $V \to (V^*)^*$ ein Isomorphismus ist, entspricht $U^\perp \subset (V^*)^*$ ein Unterraum $U' \subset V$, der unter $V \to (V^*)^*$ isomorph auf $U^\perp$ abgebildet wird.

Es ist also $\dim(U^\perp) = \dim(U') = \dim(V) - \dim(U)$.

Jetzt ist $v \in U'$ genau dann, wenn das Bild φ_v von v in $(V^*)^*$ aus $U^\perp$ ist. Es ist

$$\varphi_v(f) = f(v) \quad .$$

$\varphi_v \in U'$ genau dann, wenn $\varphi_v(f) = 0$ für alle $f \in U$.

Also ist $v \in U'$ genau dann, wenn $f(v) = 0$ ist, d.h. wenn $v \in U^\perp$ (bezüglich V). Daraus folgt (ii).

Zu (iii) Nach Definition ist für $W \subset V$ $\langle W, W^\perp \rangle = 0$, also $W \subset (W^\perp)^\perp$. Andererseits ist

$$\begin{aligned} \dim(W^\perp)^\perp &= \dim(V) - \dim(W^\perp) \\ &= \dim(V) - (\dim(V) - \dim(W)) \\ &= \dim(W) \quad . \end{aligned}$$

Daraus folgt: $W = (W^\perp)^\perp$.
Entsprechend zeigt man $(U^\perp)^\perp = U$.

10.9 Bemerkung: Ein endlich-dimensionaler Vektorraum V ist also isomorph zu $(V^*)^*$, aber natürlich auch zu V^* wegen $\dim(V) = \dim(V^*)$. Während die Isomorphie $V \cong V^*$ willkürlich gewählt werden muß, ist der Isomorphismus $V \cong (V^*)^*$ völlig natürlich ohne Festlegung einer Basis gegeben.

10.10 Funktorielles Verhalten:

a) Gegeben sei eine lineare Abbildung $\varphi : V \to W$. Diese induziert in natürlicher Weise eine lineare Abbildung, die sog. adjungierte Abbildung,

$$\left.\begin{array}{cccc} \varphi^* : & W^* & \to & V^* \\ & f & \mapsto & (f \circ \varphi) \end{array}\right\} \quad ,$$

das heißt, der Linearform $f : W \to K$ wird durch Komposition eine Linearform

$$\begin{array}{ccc} V & \xrightarrow{\varphi} W \xrightarrow{f} & K \\ V & \xrightarrow{(f \circ \varphi)} & K \end{array}$$

zugeordnet.

Man zeigt leicht: φ^* ist wieder linear (Übungen oder Praktikum).

b) Hat man $V \xrightarrow{\varphi} W \xrightarrow{\psi} U$, so erhält man per Dualität lineare Abbildungen

$$U^* \xrightarrow{\psi^*} W^* \xrightarrow{\varphi^*} V^* \quad .$$

Übung: Zu zeigen: $(\psi \circ \varphi)^* = \varphi^* \circ \psi^*$

c) Man hat einen sogenanntes kommutatives Diagramm:

$$\begin{array}{ccc} V & \xrightarrow{c_V} & (V^*)^* \\ \varphi \downarrow & & \downarrow (\varphi^*)^* \\ W & \xrightarrow{c_W} & (W^*)^* \end{array}$$

Beweis: (Übung)

11 Lineare Gleichungssysteme

Wir betrachten lineare Gleichungssysteme der Form

$$\begin{array}{ccccccccc}
a_{11}\xi_1 & + & a_{12}\xi_2 & + & \dots & + & a_{1n}\xi_n & = & b_1 \\
a_{21}\xi_1 & + & a_{22}\xi_2 & + & \dots & + & a_{2n}\xi_n & = & b_2 \\
\vdots & & & & \vdots & & & & \vdots \\
a_{m1}\xi_1 & + & a_{m2}\xi_2 & + & \dots & + & a_{mn}\xi_n & = & b_m
\end{array}$$

In kompakter Form also

$$\sum_{j=1}^{n} a_{ij}\xi_j = b_i \qquad (i = 1, \dots, m)$$

in den Unbekannten $\xi_1, \dots, \xi_n$.

Das zugehörige homogene lineare Gleichungssystem ist dann

$$\sum_{j=1}^{n} a_{ij}\xi_j = 0 \qquad (i = 1, \dots, m)$$

Wir können auch alles in Matrixform schreiben, mit

$$x = (\xi_1, \dots, \xi_n) \text{ und } b = (b_1, \dots, b_m)$$

als Spaltenvektoren und der Matrix $A = (a_{ij})_{\substack{1 \le i \le m \\ 1 \le j \le n}}$.

$$Ax = b \qquad \text{für das inhomogene System,}$$

$$Ax = 0 \qquad \text{für das homogene System.}$$

Betrachten wir $x \mapsto Ax$ als lineare Abbildung so sieht man sofort

11.1 Satz: Die Lösungsmenge des homogenen linearen Gleichungssystems

$$Ax = 0$$

hat die Struktur eines Vektorraums. Dieser ist gerade der Kern der durch A gegebenen linearen Abbildung A. Die Dimension des Lösungsraums ist gegeben durch $n - \operatorname{rang}(A)$.

Beweis: Klar.

11.2 Satz: Die Lösungsmenge des inhomogenen linearen Gleichungssystems

$$Ax = b$$

ist ein affiner Teilraum des K^n. Ist er nicht leer, so ist der zugehörige Vektorraum des affinen Lösungsraums gerade der Kern von A bzw. der Lösungsraum des zugehörigen homogenen linearen Gleichungssystems.

Beweis: Ist $x_0 \in K^n$ mit $Ax_0 = b$, $x \in \mathrm{Ker}(A)$, d.h. $Ax = 0$, so folgt:

$$A(x + x_0) = Ax + Ax_0 = b$$

Ist umgekehrt $y \in K^n$ eine Lösung der inhomogenen linearen Gleichung, so erfüllt $x := (y - x_0)$ die homogene Gleichung$Ax = b$. Daraus folgt sofort die Behauptung. □

Nützlich ist folgendes Kriterium für die Existenz einer Lösung von $Ax = b$

11.3 Satz: Bezeichnet (A, b) die mit der Spalte b erweiterte Matrix A, so gilt: $Ax = b$ hat wenigstens eine Lösung genau dann, wenn gilt $\mathrm{rang}(A) = \mathrm{rang}(A, b)$, d.h. wenn der Rang von A und der Rang der erweiterten Matrix(A,b) zusammenfallen.

Beweis: Wir betrachten etwa den Spaltenrang der Matrizen. Bezeichnet $a_1, \ldots, a_n$ die n Spalten von A in K^m, so bedeutet das Bestehen der Gleichung $Ax = b$ gerade die Relation

$$\sum_{i=1}^{n} \xi_i a_i = b$$

Also: Besteht eine Lösung $(\xi_1, \ldots, \xi_n)$, so lässt sich die Spalte b gerade aus den Spalten $a_1, \ldots, a_n$ linear kombinieren. Besteht umgekehrt eine Relation der Form

$$\sum_{i=1}^{n} \xi_i a_i = b$$

so ist $x = (\xi_1, \ldots, \xi_n)$ gerade eine Lösung. □

11.4 Praktische Lösungsmethoden: Wie löst man also gegebenenfalls ein inhomogenes Gleichungssystem der Form $Ax = b$. Wir wenden elementare Zeilenumformungen für die Matrizen A bzw. (A, b) an. Wir bringen damit A bzw. (A, b) simultan auf Zeilenstufenform. Dies bedeutet ja offenbar nichts anderes als die entsprechende Operation statt mit den Zeilen mit den Gleichungen selber durchzuführen, was offenbar die Lösungsmenge nicht ändern wird.

Wir können also nach endlich vielen Schritten erreichen, dass A und (A, b) so aussehen:

$$\xi_{i_1} + \sum_{\substack{j=i_1+1 \\ j \neq i_2, \ldots i_r}}^{n} \tilde{\alpha}_{1j} \xi_j = \tilde{b}_1$$

$$\xi_{i_2} + \sum_{\substack{j=i_2+1 \\ j\neq i_2,\dots i_r}}^{n} \tilde{\alpha}_{2j}\xi_j = \tilde{b}_2$$

$$\vdots$$

$$\xi_{i_r} + \sum_{j=i_r+1}^{n} \tilde{\alpha}_{rj}\xi_j = \tilde{b}_r$$

$$0 \qquad\qquad = \tilde{b}_{r+1}$$

$$\vdots$$

$$0 \qquad\qquad = \tilde{b}_m$$

Es existiert also genau dann eine Lösung wenn gilt

$$\tilde{b}_{r+1} = \ldots = \tilde{b}_m = 0$$

d.h. ja gerade

$$\text{Spaltenrang}(A) = \text{Spaltenrang}(A, b)$$

Wenn dies aber der Fall ist, so können wir das Gleichungssystem jetzt folgendermaßen (in Parameterform) lösen: Wir wählen die ξ_j mit $j \neq i_1, \ldots, i_r$beliebig, dann setzen wir

$$\xi_{i1} = - \sum_{\substack{j=i_1+1 \\ j\neq i_2,\dots i_r}}^{n} \tilde{\alpha}_{1j}\xi_j + \tilde{b}_1$$

$$\vdots$$

$$\xi_{ir} = - \sum_{j=ir+1}^{n} \tilde{\alpha}_{rj}\xi_j + \tilde{b}_r$$

Damit ist, die Menge aller Lösungen in parametrisierter Form gegeben. □

11.5 Der Fall n=m: Besonders wichtig ist der Fall $n = m$: n lineare Gleichungen in n Unbekannten.

$$Ax = b$$

Natürlich könnten wir das Verfahren eben anwenden. Ist aber die Matrix A invertierbar bzw. die durch A gegebene lineare Abbidung invertierbar, so ergibt sich sofort

$$A^{-1}(Ax) = A^{-1}b$$

bzw.

$$x = A^{-1}b$$

Es reicht also aus, das Inverse einer $(n \times n)$ - Matrix zu berechnen. Hierfür gibt es folgendes einfache Verfahren, das wir zunächst beschreiben und dann begründen: Wir schreiben die Matrix A und die Einheitsmatrix nebeneinander,

$$A \quad E$$

und führen an A elementare Zeilenumformungen wie gehabt so durch, dass wir A in Zeilenstufenform erreichen. Offenbar gilt: A invertierbar $\Leftrightarrow$ $\operatorname{rang}(a) = n \Leftrightarrow$ die Zeilenstufenform von A ist von der Gestalt

$$\begin{pmatrix} \alpha_{11} & & * \\ & \ddots & \\ 0 & & \alpha_{nn} \end{pmatrix}$$

dabei bedeutet 0 unterhalb der Diagonalen, dass dort alle Einträge Null sind, $*$ oberhalb der Diagonalen, dass die Einträge dort beliebig sind. Die $\alpha_{11}, \ldots, \alpha_{nn}$ sind $\neq 0$. Durch weitere Zeilenoperationen kann man jetzt weiter, beginnend mit der letzten Zeile, die man geeignet von den oberen Zeilen abzieht, erreichen, dass auch alle Einträge oberhalb der Diagonalen Null werden. Damit kommt man auf die Einheitsmatrix, wenn man in einem letzten Schritt noch auf beiden Seiten von links mit

$$\begin{pmatrix} \alpha_{11}^{-1} & & 0 \\ & \ddots & \\ 0 & & \alpha_{nn}^{-1} \end{pmatrix}$$

multipliziert. Simultan führen wir genau die gleichen elementaren Umformungen an der Einheitsmatrix E durch, so dass sich am Ende die Matrix B ergibt.

11.6 Behauptung: Es ist $B = A^{-1}$

Beweis: Jede elementare Umformung bedeutet, wie wir schon gesehen haben, die Multiplikation von links mit einer geeigneten elementaren Matrix Z_i bzw. am Ende mit

$$\begin{pmatrix} \alpha_{11}^{-1} & & 0 \\ & \ddots & \\ 0 & & \alpha_{nn}^{-1} \end{pmatrix}$$

Insgesamt etwa $(Z_r \ldots Z_2 Z_1)A{=}E$ und $(Z_r \ldots Z_2 Z_1)E = B$. Aber wegen der ersten Gleichung folgt sofort

$$Z_r Z_{r-1} \ldots Z_1 = A^{-1}$$

Daher ergibt sich $B{=}A^{-1}$. Genau das war zu zeigen. $\square$

Bemerkung: Bricht das Verfahren ab, bevor man die Einheitsmatrix erreicht hat, so bedeutet dies offenbar, dass $\operatorname{rang}(A) < n$, d.h. A gar nicht invertierbar ist. Diese zusätzliche Information wird also auch durch das Verfahren mitgeliefert. Natürlich kann, auch wenn

A nicht invertierbar bzw. äquivalent $\operatorname{rang}(A) < n$ ist, die Gleichung $Ax = b$ lösbar sein. Es sind dann die vorher entwickelten Methoden anzuwenden.

Zum Schluß dieses Kapitels geben wir noch ein etwas modifiziertes Kriterium für die Lösbarkeit linearer Gleichungssysteme an, was theoretisch sehr zufriedenstellend ist. Gegeben sei das lineare Gleichungssystem

$$\sum_{j=1}^{n} a_{ij}\xi_j = b_i \qquad (i = 1, \ldots, m)$$

Wir fassen

$$f_i(\xi_1, \ldots, \xi_n) := \sum_{j=1}^{n} a_{ij}\xi_j \qquad (i = 1, \ldots, m)$$

als Linearform auf dem K^n auf.

11.7 Satz: Das Gleichungssystem $f_i(x) = b_i \quad (i = 1, 2, \ldots, m)$ ist lösbar genau, wenn aus dem Bestehen einer Relation $\sum_{i=1}^{m} \lambda_i f_i = 0$ im Dualraum des K^n stets folgt:

$$\sum_{i=1}^{m} \lambda_i b_i = 0$$

Beweis: Existiert eine Lösung $x = (\xi_1, \ldots, \xi_n)$ des obigen Systems, so folgt aus $\sum_{i=1}^{m} \lambda_i f_i = 0$ sofort $\sum_{i=1}^{m} \lambda_i f_i(x) = 0$ und daher $\sum_{i=1}^{m} \lambda_i b_i = 0$.

Wir zeigen die Umkehrung: Wir betrachten die zur Matrix

$$A = (a_{ij})_{\substack{1\le i\le m\\ 1\le j\le n}}$$

gehörende Abbildung

$$A : K^n \to K^m$$

Wir komponieren mit der linearen Abbildung $K^m \to K, (y_1, \ldots, y_m) \mapsto \sum_{i=1}^{m} \lambda_i y_i$. Die zusammengesetzte Abbildung ist dann $\sum_{i=1}^{m} \lambda_i f_i : K^n \to K$. Wir nehmen jetzt an, es wäre $(b_1, \ldots, b_m) \notin \operatorname{Im}(A : K^n \to K^m)$. Dann gibt es sicherlich eine Linearform $\varphi : K^m \to K$, mit $\varphi(b_1, \ldots, b_m) \neq 0, \varphi| \operatorname{Im}(A) = 0$ (Übung!). φ sei dabei gegeben durch $\varphi(y_1, \ldots, y_m) = \sum_{i=1}^{m} \lambda_i y_i$. Dann ist also $\sum_{i=1}^{m} \lambda_i b_i \neq 0$, aber $\sum_{i=1}^{m} \lambda_i f_i = 0$ (wegen $\varphi = 0$ auf $\operatorname{Im}(A)$!).

Dies ist ein Widerspruch, also folgt: $(b_1, \ldots, b_m) \in \operatorname{Im}(A)$, d. h. das lineare Gleichungssystem $f_i(x) = b_i (i = 1, \ldots, m)$ hat eine Lösung $x = (\xi_1, \ldots, \xi_n)$. □

12 Determinanten

Betrachtet man ein Parallogramm in der Ebene bzw. ein Parallelepiped im Raum, so stellt man fest, daß sein Flächeninhalt bzw. Voluminen durch folgende Formeln gegeben wird:

i) Seien etwa $a_1, a_2 \in \mathbb{R}^2$ Vektoren, $a_i = \sum\limits_{j=1}^{2} \alpha_{ij} e_j$, so ergibt sich für die Fläche F des Parallelogramms P

$$\begin{aligned} P &= \{\lambda_1 a_1 + \lambda_2 a_2 \mid 0 \leq \lambda_1, \lambda_2 \leq 1\} \\ F &= |\alpha_{11}\alpha_{22} - \alpha_{21}\alpha_{12}| \end{aligned}$$

ii) Entsprechend, aber etwas komplizierter, erhält man im $\mathbb{R}^3$ bei drei gegebenen Vektoren
$a_1, a_2, a_3 \in \mathbb{R}^3$ mit $a_i = (\alpha_{i1}, \alpha_{i2}, \alpha_{i3})$ für das Volumen $V(P)$ des Parallelepiped

$$\begin{aligned} P &= \{\lambda_1 \alpha_1 + \lambda_2 \alpha_2 + \lambda_3 \alpha_3 \mid 0 \leq \lambda_1, \lambda_2, \lambda_3 \leq 1\} \\ F &= |\alpha_{11}\alpha_{22}\alpha_{33} + \alpha_{12}\alpha_{21}\alpha_{31} + \alpha_{13}\alpha_{21}\alpha_{32} - \alpha_{13}\alpha_{22}\alpha_{31} - \alpha_{12}\alpha_{21}\alpha_{33} - \alpha_{11}\alpha_{23}\alpha_{32}| \end{aligned}$$

Es liegt nahe, wenn man dies besser verstehen und wenn möglich auf n-dimensionale Räume verallgemeinern möchte, die Betragsstriche zunächst einmal wegzulassen. Die obigen Formeln werden dann erheblich angenehmer und sind offenbar linear in den Vektorvariablen a_1, a_2 bzw. a_1, a_2, a_3
D. h. es gilt etwa

$$f(a_1 + b_1, a_2, a_3) = f(a_1, a_2, a_3) + f(b_1, a_2, a_3)$$

usw.

Eine andere Eigenschaft, die man an den obigen Formeln sofort abliest, ist diese: Sind die Vektoren $\{a_1, a_2, a_3\}$, linear abhängig, so ist das obige Parallelepiped im $\mathbb{R}^3$ in Wirklichkeit kein eigentlicher geometrischer Körper sondern ist vielmehr (da in einer Ebene enthalten), ein Parallelogram. Daher ist in diesem Falle das Volumen $v(P) = 0$.
Diese Vorüberlegung in Betracht gezogen, liegt es nahe folgende Abbildung mit folgenden Eigenschaften zu betrachten:
Gesucht sind Abbildungen:

$$\left.\begin{array}{rccc} f: & \underset{\text{(n-mal)}}{V \times \ldots \times V} & \longrightarrow & K \\ & (a_1, \ldots, a_n) & \longmapsto & f(a_1, \ldots, a_n) \end{array}\right\}$$

mit folgenden Eigenschaften:

i) f ist ein sog. multilineare Abbildung, d.h. für $\lambda, \mu \in K$, $i \in \{1 \ldots, n\}$ beliebig gilt:

$$f(a_1, \ldots, \lambda a_i + \mu b_i, \ldots, a_n) = \lambda f(a_1, \ldots, a_i, \ldots, a_n + \mu f(a_1, \ldots, b_i, \ldots, a_n)$$

ii) Ist in dem n-Tupel von Vektoren $(a_1, \ldots, a_n)$ $a_i = a_j$ für ein Paar $i, j \in \{1, \ldots, n\}, i \neq j$, so ist

$$f(a_1, \ldots, a_n) = 0$$

Bemerkung: Ist, wie oben, $a_i = a_j$, so ist das System von Vektoren $\{a_1, \ldots, a_n\}$ sicher linear abhängig! Es reicht aber aus in (ii), das Verschwinden von f zunächst nur in der speziellen Situation oben zu fordern. Wir werden unten sehen, daß die allgemeine Aussage sich daraus ergeben wird.

12.1 Definition: Eine Abbildung mit den Eigenschaften (i) und (ii) oben nennt man eine *alternierende Multilinearform.*

Der Begriff „alternierend“ wird noch deutlicher durch folgende

12.2 Proposition: Es gilt:

$$f(a_1, \ldots, a_i, \ldots, a_j, \ldots, a_n) = -f(a_1, \ldots, a_j, \ldots, a_i, \ldots, a_n)$$

für $i, j \in \{1, \ldots, n\}$, d. h. vertauscht man die Argumente a_i, a_j, so ergibt sich der negative Wert der Multilinearform.

Beweis: Wie der Beweis gleich zeigt, kann o.B.d.A. angenommen werden $i = 1, j = 2$. Wir betrachten:

$$\begin{aligned} 0 &= f(a_1 + a_2, a_1 + a_2, a_3, \ldots, a_n) \\ &= f(a_1, a_1, a_3, \ldots, a_n) + f(a_2, a_1, a_3, \ldots, a_n) \\ &+ f(a_1, a_2, a_3, \ldots, a_n) + f(a_2, a_2, a_3, \ldots, a_n) \\ &= f(a_1, a_2, \ldots, a_n) + f(a_2, a_1, a_3, \ldots, a_n) \end{aligned}$$

□

Als letzte Bedingung stellen wir noch die folgende Forderung:

iii) $f(e_1, e_2, \ldots, e_n) = 1$, falls $\{e_1, \ldots, e_n\}$ eine festgelegte Standardbasis ist.

Bemerkung: Ist $V = K^n$ der Standardvektorraum, so ist $\{e_1, \ldots, e_n\}$ etwa die Standardbasis.

12.3 Proposition: Ist $\{e_1, \ldots, e_n\}$ irgendeine Basis des Vektorraums V,

$$f : V \times \ldots \times V \longrightarrow K$$

eine Multilinearform,

$$a_i = \sum_{j=1}^{n} \alpha_{ij} e_j \ (i = 1, \ldots, n)$$

ein System von Vektoren, so gilt:

$$f(a_1, \ldots, a_n) = \sum_{\pi} \alpha_{1,\pi(1)}, \ldots, \alpha_{1,\pi(n)} f(e_{\pi(1)}, \ldots, e_{\pi(n)})$$

dabei versteht sich die Summe über die Menge aller Permutationen (= bijektive Abbildung $\{1, \ldots, n\} \longrightarrow \{1, \ldots, n\}$) von $\{1, \ldots, n\}$.

Beweis: Man verwendet die Linearitätseigenschaft und berücksichtigt, daß Terme, in denen zwei Argumente gleich sind, Null sind. □

Natürlich kann man die Terme

$$f(e_{\pi(1)}, \ldots, e_{\pi(n)})$$

durch geeignete Vertauschung alle auf die Form

$$\pm f(e_1, e_2, \ldots, e_n)$$

bringen. Was nicht klar ist, ob das Vorzeichen dabei nicht, je nach dem, wie man die Vertauschungen durchführt, zu wechselnden Vorzeichen führen würde. Das dies tatsächlich *nicht* der Fall ist, ist überhaupt nicht trivial.

12.4 Definition: Sei $\pi \in S_n$ eine Permutation von $\{1, \ldots, n\}$. Es bezeichne $\alpha(\pi)$ die Anzahl der Paare $i, j \in \{1, \ldots, n\}$ mit $i < j$ und $\pi(i) > \pi(j)$, der sog. *Fehlstellungen.*

Es sei $\operatorname{sgn}(\pi) = (-1)^{\alpha(\pi)}$. Es gilt dann Folgendes:

12.5 Satz: Die Anzahl der Vertauschungen von Ziffern, mit denen man aus einer Permutation $\{\pi(1), \ldots, \pi(n)\}$ die Identität herstellen kann, ist gerade, wenn $\operatorname{sgn}(\pi) = +1$ ist, ungerade, wenn $\operatorname{sgn}(\pi) = -1$ ist. Insbesondere ist sie stets entweder gerade oder ungerade.

Den Beweis dieses Resultates stellen wir kurze Zeit zurück, um mit der Theorie der Determinanten weiterzumachen. Wir erhalten jetzt also

$$f(a_1, \ldots, a_n) = \sum_{\pi \in S_n} \operatorname{sgn}(\pi) \alpha_{1,\pi(1)}, \ldots, \alpha_{n,\pi(n)} \cdot f(e_1, \ldots, e_n)$$

Wir sehen also: Alternierende Multilinearformen sind bestimmt durch Festlegung des Wertes $f(e_1, \ldots, e_n)$ einer Basis $\{e_1, \ldots, e_n\}$ von V.

12.6 Satz: Durch den Ausdruck

$$f(a_1,\dots,a_n) = \sum_{\pi\in S_n} \operatorname{sgn}(\pi)\alpha_{1,\pi(1)},\dots,\alpha_{n,\pi(n)} \cdot f(e_1,\dots,e_n)$$

mit

$$a_i = \sum_{j=1}^{n} \alpha_{ij}e_j \ (i = 1\dots n)$$

wird eine alternierende Multilinearform festgelegt, die trivial ist, genau dann, wenn $f(e_1,\dots,e_n) = 0$ ist.

Beweis: Die Multilinearität, d.h. die Linearität in jeder der Variablen $a_1, a_2, \dots$ entnimmt man sofort der Formel. Ist etwa $a_i = a_j$, so schreibe man

$$\begin{aligned} f(a_1,\dots,a_n) = \sum_{\substack{\pi\in S_n\\ \pi(i)<\pi(j)}} &(\alpha_{1,\pi(1)}\dots\alpha_{i,\pi(i)}\dots\alpha_{j,\pi(j)}\dots\alpha_{n,\pi(n)}\operatorname{sgn}(\pi) \\ &+\alpha_{1,\pi(1)}\dots\alpha_{j,\pi(j)}\dots\alpha_{i,\pi(i)}\dots,\alpha_{n,\pi(n)}\operatorname{sgn}(\tilde\pi)) \quad , \end{aligned}$$

dabei entsteht die Permutation $\tilde\pi$ aus π jeweils durch Vertauschen von $\pi(i)$ und $\pi(j)$. Dann gilt aber $\operatorname{sgn}(\tilde\pi) = (-1)\operatorname{sgn}(\pi)$. Das sieht man so ein. Vertauscht man $\pi(i), \pi(i+1)$, so ändert sich die Anzahl der Fehlstellung um $+1$ oder -1, in jedem Fall ergibt sich für die so geänderte Permutation das Signum $(-1)\operatorname{sgn}(\pi)$. Da nun eine Vertauschung von $\pi(i)$ und $\pi(j)$ durch die Folge von Vertauschung von $\pi(i)$ mit $\pi(i+1),\dots,\pi(j)$, also $(j-i)$ Vertauschungen und anschließende Vertauschung von $\pi(j)$ mit $\pi(j-1),\dots,\pi(i+1)$, also von $(j-i+1)$ Vertauschungen bewerkstelligt wird, so ist die Anzahl der Vertauschungen insgesamt $2(j-i)+1$, also ungerade, woraus folgt:

$$\operatorname{sgn}(\tilde\pi) = (-1)^{2(j-i)+1}\operatorname{sgn}(\pi) = (-1)\operatorname{sgn}(\pi)$$

Daraus folgt aber sofort

$$f(a_1,\dots,a_i,\dots,a_j,\dots,a_n) = 0$$

falls $a_i = a_j$.
Ist jetzt $f(e_1,\dots,e_n) = 1$ etwa, so ist offenbar f nicht trivial. □

12.7 Bemerkung: Man überlegt sich leicht, daß man die alternierenden Multilinearformen zu einem Vektorraum macht, indem wie üblich gesetzt wird:

$$\begin{aligned} (f+g)(a_1,\dots,a_n) &= f(a_1,\dots,a_n) + g(a_1,\dots,a_n) \\ (\lambda f)(a_1,\dots,a_n) &= \lambda f(a_1,\dots,a_n) \end{aligned}$$

Die obige Überlegung zeigt also: Der Vektorraum der alternierenden Multilinearformen ist *eindimensional.*

Determinanten von Endomorphismen

Sei jetzt $\varphi : V \to V$ ein beliebiger Endomorphismus des Vektorraums V, $f : V \times \ldots \times V \to K$ eine beliebige nichttriviale alternierende Multilinearform. Betrachte die Abbildung

$$\tilde{f} : V \times \ldots \times V \longrightarrow K \quad ,$$

gegeben durch

$$\tilde{f}(a_1, \ldots, a_n) := f(\varphi(a_1), \ldots, \varphi(a_n))$$

Man sieht sofort: Auch $\tilde{f}$ ist eine alternierende Multilinearform auf V. Daher ist sie proportional zur Form f, d.h. es gibt ein Element $\det(\varphi) \in K$, $\tilde{f} = \det(\varphi) \cdot f$.

12.8 Definition und Satz: Es gibt ein eindeutig bestimmtes Element aus K, die sog. Determinante $\det(\varphi)$ des Endomorphismus von φ, so daß gilt:

$$f(\varphi(a_1), \ldots, \varphi(a_n)) = \det(\varphi) f(a_1, \ldots, a_n)$$

für $a_1, \ldots, a_n \in V$.

12.9 Bemerkung: Man rechnet sofort nach, daß $\det(\varphi)$ unabhängig von der Wahl der alternierenden Form f ist, da f bis auf einen skalaren Faktor eindeutig bestimmt ist.

12.10 Satz:

i) Sind $\varphi, \psi \in \mathrm{End}(V)$, so gilt $\det(\varphi \circ \psi) = \det(\varphi) \cdot \det(\psi)$.

ii) $\varphi : V \to V$ ist ein Isomorphismus, genau dann, wenn $\det(\varphi) \neq 0$ ist.

iii) Es gilt $\det(id) = 1$ für die Idendität $id \in \mathrm{End}(V)$.
Es gilt $\det(\varphi^{-1}) = \det(\varphi)^{-1}$, falls φ^{-1} existiert.

Beweis: Es ist offenbar einerseits

$$f((\varphi \circ \psi)(e_1), \ldots, (\varphi \circ \psi)(e_n)) = \det(\varphi \circ \psi) f(e_1, \ldots, e_n).$$

Andererseits ist

$$\begin{aligned} & f((\varphi \circ \psi)(e_1), \ldots, (\varphi \circ \psi)(e_n)) \\ = \; & f(\varphi(\psi(e_1)), \ldots, \varphi(\psi(e_n)) \\ = \; & \det(\varphi) \; f(\psi(e_1), \ldots, \psi(e_n)) \\ = \; & \det(\varphi) \; \det(\psi) \; f(e_1, \ldots, e_n) \end{aligned}$$

Damit folgt (i); (ii) und (iii) sind einfache Übungen.

12.11 Bemerkung: Sei der Endomorphismus φ bzgl. irgendeiner Basis $\{e_1, \ldots, e_n\}$, in Dreiecksform, d. h. es ist

$$\varphi(e_i) = \alpha_{ii} e_i + \sum_{j=1}^{i-1} \alpha_{ji} e_j \quad (i = 1 \ldots n)$$

Dann ist

$$\det(\varphi) = \prod_{i=1}^{n} \alpha_{ii}$$

Beweis:

$$\begin{aligned} & f(\varphi(e_1), \ldots, \varphi(e_n)) \\ = \; & f(\alpha_{11} e_1, \alpha_{12} e_1 + \alpha_{22} e_2, \ldots, \alpha_{1n} e_1 + \ldots + \alpha_{nn} e_n) \\ = \; & (\alpha_{11} \alpha_{22} \ldots \alpha_{nn}) \; f(e_1, \ldots, e_n) \end{aligned}$$

□

Jetzt lassen sich auch leicht die Determinaten von Matrizen einführen.

Sei $A = (\alpha_{ij})_{i \leq i, j \leq n}$ eine Matrix aus $\mathbb{M}(n; K)$, $\{e_1, \ldots, e_n\}$ die Standardbasis von K^n. Wir können den durch A gegebenen Endomorphismus von V bzgl. der Basis $\{e_1, \ldots, e_n\}$ betrachten. Bezeichnet

$$\tilde{a_j} = (\alpha_{1j}, \ldots, \alpha_{nj}) \; (j = 1 \ldots n)$$

die j-te Spalte der Matrix A, so ergibt sich offenbar nach Definition

$$\begin{aligned} f(\tilde{a_1}, \ldots, \tilde{a_n}) &= \det(A) \cdot f(e_1, \ldots, e_n) \\ &= \det(A), \text{ wenn } f(e_1, \ldots, e_n) = 1 \text{ ist.} \end{aligned}$$

Wir zeigen nun zuerst:

12.12 Satz: Für Matrizen $A \in \mathbb{M}(n; K)$ gilt:

$$\det(A) = \det(A^t)$$

Beweis: Wir schreiben die Entwicklungsformel für beide Ausdrücke hin, wobei wir noch benutzen:

$$\det(A) = f(\tilde{a_1}, \ldots, \tilde{a_n})$$
$$\det(A^t) = f(a_1, \ldots, a_n)$$

Dabei natürlich $a_i = (\alpha_{i1}, \ldots, \alpha_{in})$ die i-te Zeile für $1 \leq i \leq n$. Es ist dann:

$$\begin{aligned}
\det(A) &= \sum_{\pi \in S_n} \operatorname{sgn}(\pi)\ \alpha_{1,\pi(1)} \ldots \alpha_{n,\pi(n)} \\
\det(A^t) &= \sum_{\pi \in S_n} \operatorname{sgn}(\pi)\ \tilde{\alpha}_{1,\pi(1)} \ldots \tilde{\alpha}_{n,\pi(n)} \\
&= \sum_{\pi \in S_n} \operatorname{sgn}(\pi)\ \alpha_{\pi(1),1} \ldots \alpha_{\pi(n),n} \\
&= \sum_{\pi \in S_n} \operatorname{sgn}(\pi)\ \alpha_{1,\pi^{-1}(1)} \ldots \alpha_{n,\pi^{-1}(n)}
\end{aligned}$$

(hierbei hat man beim letzten Schritt die Terme nur umgeordnet, man beachte $\alpha_{\pi(i),i} = \alpha_{j,\pi^{-1}(j)}$ mit $j = \pi(i)$)

Offenbar gilt aber:

$$\operatorname{sgn}(\pi^{-1}) = \operatorname{sgn}(\pi) \quad ,$$

denn ist etwa $1 \leq i < j \leq n$ und $\pi(i) > \pi(j)$ eine Fehlstellung von π, so folgt wegen $\pi^{-1}(\pi(i)) = i < \pi^{-1}(\pi(j)) = j$, aber $\pi(i) > \pi(j)$), sofort, daß $(\pi(j), \pi(i))$ eine Fehlstellung von π^{-1} ist. So folgt:

$$\det(A^t) = \sum_{\pi \in S_n} \operatorname{sgn}(\pi^{-1})(\alpha_{1,\pi^{-1}(1)} \ldots \alpha_{n,\pi^{-1}(n)})$$

Mit π durchläuft aber auch die inverse Permutation π^{-1} die ganze Menge S_n. Daher können wir, statt über π, auch über π^{-1} summieren und anschließend π durch π^{-1} ersetzen. Dann folgt:

$$\det(A) = \det(A^t)$$

□

12.13 Folgerung: Für $A \in \mathbb{M}(n; K)$ ist $\det(A)$ (als Determinante des Endomorphismus A bzgl. der Basis $\{e_1, \ldots, e_n\}) = f(a_1, \ldots, a_n)$.

12.14 Folgerung: Sei $A \in \mathbb{M}(n; K)$, so gilt:

i) Vertauscht man in A zwei Spalten, so ändert sich $\det(A)$ um den Faktor (-1).

ii) Die Spalten von A sind linear abhängig, genau dann, wenn $\det(A) = 0$ ist.

iii) Multipliziert man eine Spalte der Matrix A mit einen Faktor $\lambda \in K$, so multipliziert sich die Determinante mit λ.

iv) Addiert man das λ-fache einer Spalte von A zu einer anderen Spalte von A und lässt die Matrix ansonsten unverändert, so ändert sich der Wert $\det(A)$ nicht.

Beweis: Folgt sofort aus $\det(A) = \det(A^t)$ und den entsprechenden Zeileneigenschaften.

□

12.15 Satz: Seien $A, B \in \mathbb{M}(n; K)$ $n \times n$ -Matrizen. Dann gilt

$$\det(AB) = \det(A)\ \det(B)$$

Ist A eine invertierbare Matrix, so gilt:

$$\det(A^{-1}) = \det(A)^{-1}$$

Speziell ist A invertierbar genau dann, wenn $\det(A) \neq 0$ ist.

Beweis: Wir fassen $A, B : K^n \to K^n$ als Endomorphismen auf. Alle Aussagen folgen aus den entsprechenden Aussagen in **12.10** über Determinanten von Endomorphismen. $\square$

12.16 Definition: Sei $A = (\alpha_{ij})_{1 \leq i,j \leq n} \in \mathbb{M}(n; K)$ eine $(n \times n)$ -Matrix. Seien $i, j \in \{1, \ldots, n\}$. Die zu α_{ij} komplementäre Matrix A_{ij} erhält man aus A durch Streichen der i-ten Zeile und der j-ten Spalte.

12.17 Satz: (*Laplace'sche Entwicklungsformel*) Es gilt:

$$\det(A) = \sum_{j=1}^{n} (-1)^{i+j} \alpha_{ij}\ \det(A_{ij})$$

dabei $i \in \{1, \ldots, n\}$ eine feste Zahl.

12.18 Bemerkung: Dies ist die sog. Entwicklung nach Zeilen, entsprechend kann man nach einer Spalte entwickeln, dann ergibt sich die Formel

$$\det(A) = \sum_{i=1}^{n} (-1)^{i+j} \alpha_{ij}\ \det(A_{ij})$$

dabei $j \in \{1, \ldots, n\}$.

Beweis: Wir betrachten beide Seiten als Funktion ihrer Zeilenvektoren, im übrigen führen wir einen Induktionsbeweis. Die Fälle $n = 1, n = 2$ sind klar.
Induktionsschritt: Beide Seiten sind linear in den Zeilen. Ist A die Einheitsmatrix, so ist die Formel richtig. Seien endlich zwei Zeilen gleich, etwa die die k-te und l-te. Also

$$\left.\begin{array}{rcl} a_k & = & (\alpha_{k1}, \ldots, \alpha_{kn}) \\ = \quad a_l & = & (\alpha_{l1}, \ldots, \alpha_{ln}) \end{array}\right\}$$

Sei zunächst $k \neq i, l \neq i$. Dann folgt sofort, daß beide Seiten gleich sind. Sei jetzt etwa $k = i, l \neq i$. In diesem Falle entwickelt man auf der rechten Seite nach Induktionsannahme die Determinanten A_{ij} nach der letzen Zeile (bzgl. A). Man sieht dann leicht, dass auch die rechte Seite Null wird.
Damit erfüllen beide Seiten in **12.17** die kennzeichnenden Eigenschaften einer Determinante. Die Behauptung folgt. $\square$

Eine unmittelbare Folgerung aus der Laplaceschen Entwicklungsformel ist folgende Formel für das Inverse einer Matrix:

12.19 Satz: Sei $A = (\alpha_{ij})_{1\leq i,j\leq n}$ eine $n \times n$-Matrix mit $\det(A) \neq 0$. Dann ist A invertierbar und es ist

$$A^{-1} = \frac{1}{\det(A)}(B_{ij})_{1\leq i,j\leq n}$$

mit

$$B_{ij} = (-1)^{i+j}\ \det(A_{ji})(i \leq i, j \leq n)$$

Beweis: Folgt sofort aus **12.17** Durch Ausmultiplizieren von $A \cdot A^{-1}$. □

12.20 Folgerung: (*Cramersche Regel*) Sei

$$Ax = \underline{b}$$

ein Gleichungssystem von n linearen Gleichungen mit n Unbekannten, dabei $x = (\xi, \ldots, \xi_n)$, $\underline{b} = (\beta_1, \ldots, \beta_n)$ geschrieben als Spaltenvektoren. Dann gilt:
Ist $\det(A) \neq 0$, also A invertierbar, so ist

$$\xi_i = \frac{1}{\det(A)}\ \det(\tilde{A}_i) \quad \text{für } i = 1, 2, \ldots, n$$

Dabei entsteht $\tilde{A}_i$ aus A durch Ersetzen der i-ten Spalte von A durch die Spalte

$$\underline{b} = \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_n \end{pmatrix}$$

Beweis: $\det(A) \neq 0 \Rightarrow A$ ist invertierbar.

$$\begin{aligned}
\underline{x} &= \begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix} = A^{-1}\underline{b} = A^{-1}\begin{pmatrix} \beta_1 \\ \vdots \\ \beta_n \end{pmatrix} \\
A^{-1} &= \frac{1}{\det(A)}\ \big((-1)^{i+j}\ \det(A_{ji})\big) \qquad \text{(nach } \mathbf{12.19}\text{)} \\
\xi_i &= \frac{1}{\det(A)}\ \sum_{j=1}^{n}(-1)^{i+j}\ \det(A_{ji})\beta_j \\
&= \frac{1}{\det(A)}\ \det(\tilde{A}_i) \quad ,
\end{aligned}$$

dabei entsteht $\tilde{A}_i$ durch Ersetzen der i-ten Spalte in A durch die Spalte

$$\begin{pmatrix} \beta_1 \\ \vdots \\ \beta_n \end{pmatrix} = \underline{b}$$

□

Es sei hier vermerkt, daß die Entwicklungsformeln für die Determinanten zur praktischen Berechnung recht ungeeignet sind. So hätte man bei einer Berechnung eine 8×8-Matrix bereits $8! = 40320$ Terme zu berücksichtigen. Das Zeilentauschverfahren ist viel einfacher!

Zum Schluss noch eine häufig nützliche Formel:

12.21 Satz: Sei die Matrix

$$X = \begin{pmatrix} A & B \\ 0 & C \end{pmatrix} \in \mathbb{M}(n; K)$$

ferner $A \in \mathbb{M}(r; K), C \in \mathbb{M}(n-r; K)$, d. h. X ist in Blockform gegeben. Dann gilt:

$$\det(X) = \det(A)\ \det(C) \quad .$$

Beweis: Es ist $\det(C) = 0$ genau dann, wenn die Zeilenvektoren von C linear abhängig sind. Dann sind aber die entsprechenden durch Null verlängerten Zeilen von X ebenfalls linear abhängig. Damit wäre $\det(X) = 0$, ebenso $\det(C) = 0$ und die Aussage des Satzes richtig. Seien also die Zeilen von C linear unabhängig! Durch Zeilenoperationen können wir dann B zu Null machen, ohne daß der Wert von $\det(X)$ geändert wird. Sei dann $B = 0$. Dann fassen wir beide Seiten als Multilinearformen

$$f_i(a_1, \ldots, a_r; c_1, \ldots, c_{n-r}) \quad (i = 1, 2)$$

in den Zeilen von A bzw. C auf.
Man stellt leicht fest:

i) Beide Seiten sind multilinear in den $(a_1, \ldots, a_r; c_1, \ldots, c_{n-r})$

ii) Sind zwei der Zeilenvektoren in $\{a_1, \ldots, a_r\}$ bzw. $\{c_1, \ldots, c_{n-r}\}$ gleich, so sind beide Seiten Null.

iii) Ist $X = E_n$ die Einheitsmatrix, so ist $A = E_r, C = E_{n-r}$ und die Behauptung gilt.

Jetzt schließt man leicht, daß beide Seiten gleich sind (siehe auch in den Übungen). $\square$

13 Gruppen und Permutationen

Wir tragen zunächst den Beweis von Satz **12.5** nach:

Beweis: Es genügt offenbar zu zeigen: Ist $\pi \in S_n$ eine Permutation, $\tau \in S_n$ eine Transposition, so gilt für die Anzahl der Fehlstellungen:

$$\alpha(\pi) - \alpha(\pi \circ \tau)$$

ist eine ungerade Zahl. Daraus folgt dann sofort:

$$\operatorname{sgn}(\pi) = \operatorname{sgn}(\pi \circ \tau)(-1)$$

Man sieht sofort, daß bei Vertauschung etwa der Terme $\pi(i)$,$\pi(i+1)$ in der Permutation $\{\pi(1), \ldots, \pi(n)\}$ sich die Anzahl $\alpha(\pi)$ der Fehlstellungen genau um ± 1 ändert. Die neue Permutation kann geschrieben werden als $\pi \circ (i, i+1)$, dabei bezeichnet $(i, i+1)$ die Permutation von $\{1, \ldots, n\}$, die nur i und $i+1$ vertauscht.

13.1 Lemma: Eine Transposition $\tau = (ij)$, die i und j vertauscht, $i, j \in \{1, 2, \ldots, n\}$, ist ein Produkt einer ungeraden Zahl von Transpositionen der Form $(\zeta, \zeta + 1)$.

Beweis: Schreibe (mit $i < j$)

$$(i, j) = [((j-1, j) \cdot \cdots \cdot (i+2, i+3)(i+1, i+2)(i, i+1) \ldots (j-1, j-2)(j-1, j))]$$

Dies sind gerade $(j - (i+1) + 1) + (j - 1 - (i+1) + 1) = (j - i) + (j - i) - 1$ Terme, also eine ungerade Zahl. □

Hieraus folgt jetzt:

13.2 Satz: Für alle Darstellungen von $\pi \in S_n$ als Produkt von Transpositionen ist die Anzahl der Faktoren entweder durchweg gerade oder durchweg ungerade.

Beweis: Ist zunächst τ eine einzelne Transposition, so gilt für die Fehlstellungen: $\alpha(\pi\tau) - \alpha(\pi)$ ist eine ungerade Zahl. Hinzu schreiben wir τ als Produkt von Transpositionen $(j, j+1)$ und verwenden die obigen Resultate.

Dann folgt aber sofort: Ist $\pi = \tau_1 \ldots \tau_r$, so folgt $\alpha(\pi\tau_r \ldots \tau_1) - \alpha(\pi) = \alpha(\pi\tau_r) - \alpha(\pi) + \alpha(\pi\tau_r\tau_{r-1}) - \alpha(\pi\tau_r) + \cdots + (\alpha(\pi\tau_r \ldots \tau_1) - \alpha(\pi\tau_r \ldots \tau_2))$.

Als Summe von r ungeraden Zahlen ist dies gerade bzw. ungerade genau, wenn die Anzahl r der Faktoren gerade oder ungerade ist. Aber es ist $\pi\tau_1 \ldots \tau_r = 1$ die identische Permutation und $\alpha(1) = 0$. Also folgt: $\alpha(\pi)$ ist gerade oder ungerade, je nachdem ob r gerade oder ungerade ist. Hiermit folgt die Behauptung. □

13.3 Satz: Es ist

i) $\operatorname{sgn}(\pi_1, \pi_2) = \operatorname{sgn}(\pi_1)\operatorname{sgn}(\pi_2)$

ii) $\operatorname{sgn}(\pi^{-1}) = \operatorname{sgn}(\pi)$

iii) $\operatorname{sgn}(1) = 1$

Beweis: Seien $\pi_1 = \tau_1 \dots \tau_r$, $\pi_2 = \tau_1' \dots \tau_s'$ Produkte von Transpositionen. Dann folgt $\pi_1\pi_2 = \tau_1 \dots \tau_r \tau_1' \dots \tau_s'$. Es folgt $\operatorname{sgn}(\pi_1) = (-1)^r$, $\operatorname{sgn}(\pi_2) = (-1)^s$, $\operatorname{sgn}(\pi_1\pi_2) = (-1)^{r+s}$. Damit folgt (i). (ii) folgt zum Beispiel so: Ist $\pi = \tau_1 \dots \tau_r$ eine Darstellung von π als Produkt von Transpositionen. Dann folgt: $\pi^{-1} = \tau_r \dots \tau_1$! Also gilt $\operatorname{sgn}(\pi) = (-1)^r$ und $\operatorname{sgn}(\pi^{-1}) = (-1)^r$. Die Behauptung folgt. □

Bemerkung: Ein anderer Beweis zu (ii): Es ist $\operatorname{sgn}(1) = \operatorname{sgn}(\pi\pi^{-1}) = \operatorname{sgn}(\pi)\operatorname{sgn}(\pi^{-1})$. Hieraus folgt: $\operatorname{sgn}(\pi) = \operatorname{sgn}(\pi^{-1})$

Sehr einfach kann man den gewonnenen Sachverhalt so ausdrücken:

$$\operatorname{sgn} : S_n \to \{\pm 1\}$$

ist ein Gruppenhomomorphismus der symmetrischen Gruppe von n Objekten auf die Gruppe $\{\pm 1, \cdot\}$. Hierzu jetzt einige einfache Grundbegriffe aus der Gruppentheorie. Ausführlicher gehen wir darauf im zweiten Teil der Vorlesung ein.

13.4 Definition: Eine Gruppe $(G, \cdot)$ ist eine Menge versehen mit einer Verknüpfungsabbildung

$$\text{"}\cdot\text{"} : G \times G \to G, (x, y) \mapsto x \cdot y \quad ,$$

für die die folgenden Axiome gelten:

i) die Verknüpfung ist assoziativ, d. h. für $x, y, z \in G$ gilt: $(xy)z = x(yz)$

ii) die Gruppe besitzt ein Einselement $e \in G$, für das gilt: Es ist $x \cdot e = e \cdot x = x$ für alle $x \in G$.

iii) zu jedem Element $x \in G$ gibt es ein zu x inverses Element $x^{-1} \in G$, so daß gilt:

$$x \cdot x^{-1} = x^{-1} \cdot x = e$$

13.5 Bemerkungen: Es reicht übrigens in (ii) und (iii) jeweils nur die Existenz einer Linkseins bzw. eines Linksinversen (oder entsprechend einer Rechtseins und eines Rechtsinversen) zu fordern: Dann folgt automatisch, daß e auch $xe = x$ erfüllt, d. h. Rechtsseinelement ist bzw. daß $xx^{-1} = e$ gilt, d. h. daß x^{-1} auch Rechtsinverses zu x ist.

13.6 Folgerungen:

i) Das Einselement $e \in G$ ist eindeutig bestimmt.

ii) Das inverse Element x^{-1} ist eindeutig bestimmt.

iii) Für $x, y \in G$ beliebig gilt $(x^{-1})^{-1} = x$ und $(xy)^{-1} = y^{-1}x^{-1}$. Weiter ist $(e)^{-1} = e$.

iv) Für $a, b \in G$ ist die Gleichung $ax = b$ (bzw. $xa = b$) eindeutig lösbar und zwar durch $x = a^{-1}b$ (bzw. $x = ba^{-1}$).

Beweis:

i) Sei $e' \in G$ ein weiteres Einselement. Dann ist $e \cdot e' = e'$ (da e Linkseins ist), aber auch $e \cdot e' = e$ (da e' Rechtseins ist). Daher folgt: $e = e'$.

ii) Sei x' ein weiteres Element mit $xx' = x'x = e$ für gegebenes festes x. Wir betrachten $(x^{-1}xx')$. Mittels des Assoziativgesetzes ergibt sich einerseits $(x^{-1}x)x' = e \cdot x' = x'$, andererseits $x^{-1}(xx') = x^{-1}e = x^{-1}$. Daher zusammen $x^{-1} = x'$. Daher ist das inverse Element x^{-1} zu einem Element x eindeutig bestimmt.

iii) Wegen $xx^{-1} = x^{-1}x = e$ und der Eindeutigkeit des inversen Elements folgt sofort $(x^{-1})^{-1} = x$. Weiter ist für $x, y \in G$

$$(xy)(y^{-1}x^{-1}) = x(y(y^{-1}x^{-1})) = x((yy^{-1})x^{-1}) = x(ex^{-1}) = xx^{-1} = e$$

und ganz entsprechend

$$(y^{-1}x^{-1})(xy) = e$$

Da nach (ii) das inverse Element eindeutig bestimmt ist, folgt $(xy)^{-1} = y^{-1}x^{-1}$. Schließlich gilt $e \cdot e = e \cdot e = e$ und daher ist $e = (e)^{-1}$. □

13.7 Beispiele von Gruppen:

i) Wir hatten bereits die Gruppe der Permutationen von n Objekten, etwa der Zahlen $\{1, 2, \ldots, n\}$, die sogenannte S_n. Allgemeiner hat man die Gruppe

$$Bij(X) = \{\varphi : X \to X \mid \varphi \text{ ist bijektive Abbildung}\}$$

Diese Menge wird mit der Hintereinanderausführung von Abbildungen als Multiplikation zu einer Gruppe. Also $(\varphi \cdot \psi) := \varphi \circ \psi$. Damit ist das Einselement gerade die identische Abbildung

$$id : X \to X, x \mapsto x$$

Das Inverse eines Elementes $\varphi \in Bij(X)$ ist die inverse Abbildung $\varphi^{-1} : X \to X$.

ii) Sei V ein beliebiger K-Vektorraum. Dann ist

$$\mathrm{GL}(V) = \{\varphi : V \to V \mid \varphi \text{ ist bijektive lineare Abbildung}\}$$

mit der Komposition als Multiplikation eine Gruppe.

iii) Jeder Vektorraum V ist eine Gruppe mit der Vektoraddition als Verknüpfung. Natürlich schreibt man diese dann additiv, also $x + y$ (statt $x \cdot y$) für $x, y \in V$. Es handelt sich um eine sogenannte kommutative oder auch abelsche Gruppe (nach N. H. Abel, 1802–1829), in der $x + y = y + x$ für alle $x, y \in G$ gilt.

iv) Ist $(K, +, \cdot)$ ein beliebiger Körper, so sind $(K, +)$ und $(K^\times, \cdot)$ mit $K^\times := K \setminus \{0\}$ beide Gruppen, die additive und die multiplikative Gruppe des Körpers.

13.8 Definition: Sei G eine Gruppe, $H \subset G$ eine Teilmenge. H ist eine Untergruppe von G :$\Leftrightarrow$

i) Es ist $e \in H$

ii) Mit $x, y \in H$ ist auch $xy^{-1} \in H$.

13.9 Satz: Eine Untergruppe H von G ist insbesondere mit der von G durch Einschränkung definierten Operation selber eine Gruppe.

Beweis: Ist $x \in H$, so ist wegen $e \in H$ auch $e \cdot x^{-1} = x^{-1} \in H$. Sind $x, y \in H$, so ist auch $y^{-1} \in H$ nach der Überlegung oben und daher auch $x(y^{-1})^{-1} = xy \in H$. Da danach die Multiplikation auf H wohldefiniert ist und $e \in H$ ist, auch weiter mit $x \in H$ $x^{-1} \in H$ ist, ergeben sich die Eigenschaften (i), (ii), (iii) einer Gruppe sofort. $\square$

13.10 Beispiele:

i) Sei S_n die Permutationsgruppe aller Permutationen der Ziffern $\{1, 2, \dots, n\}$,

$$A_n := \{\pi \in S_n \mid \mathrm{sgn}(\pi) = 1\}$$

die Teilmenge aller geraden Permutationen. Offenbar ist A_n eine Untergruppe von S_n, die sogenannte alternierende Gruppe.

ii) Sei $\mathrm{GL}(V)$ die allgemeine linearen Gruppe der bijektiven linearen Abbildung des Vektorraumes V in sich. Sei

$$\mathrm{SL}(V) := \{\varphi \in \mathrm{GL}(V) \mid \det(\varphi) = 1\} \quad .$$

Offenbar ist $\mathrm{SL}(V)$ eine Untergruppe der $\mathrm{GL}(V)$, die sogenannte spezielle lineare Gruppe.

iii) Sei

$$B := \{\varphi \in \mathrm{GL}(V) \mid \varphi(e_i) \in Ke_1 + \cdots + Ke_i, \quad i = 1, \ldots, n\} \quad ,$$

dabei $\{e_1, \ldots, e_n\}$ eine Basis von V. B ist eine Untergruppe der $\mathrm{GL}(V)$.

iv) Sei $(V, +)$ ein Vektorraum, aufgefaßt als Gruppe bezüglich der Addition als Gruppenmultiplikation. Sei $W \subset V$ ein Untervektorraum. Dann ist $(W, +)$ eine Untergruppe von V.

13.11 Definition: G, G' seien Gruppen, $\varphi : G \to G'$ eine Abbildung. φ ist ein Gruppenhomomorphismus $:\Leftrightarrow$ Für alle $x, y \in G$ gilt $\varphi(xy) = \varphi(x)\varphi(y)$.

13.12 Folgerungen:

i) Es ist $\varphi(e) = e'$ für $e \in G$, $e' \in G'$ die Einselemente.

ii) Es ist $\varphi(x^{-1}) = \varphi(x)^{-1}$ für alle $x \in G$.

iii) Das Bild $\mathrm{Im}(\varphi) \subset G'$ ist eine Untergruppe.

iv) Der sogenannte Kern des Homomorphismus φ,

$$\mathrm{Ker}(\varphi) := \{x \in G \mid \varphi(x) = e'\}$$

ist eine Untergruppe und sogar ein sogenannter Normalteiler von G.

Beweis:

i) Es ist $\varphi(e \cdot e) = \varphi(e) = \varphi(e) \cdot \varphi(e)$ Daraus folgt sofort: $\varphi(e) = e'$.

ii) Offenbar ist $\varphi(xx^{-1}) = \varphi(e) = e'$, andererseits
$\varphi(xx^{-1}) = \varphi(x)\varphi(x^{-1}) = e'$ bzw.
$\varphi(x^{-1}x) = \varphi(x^{-1})\varphi(x) = e'$,
daher $\varphi(x^{-1}) = \varphi(x)^{-1}$ wegen der Eindeutigkeit des Inversen eines Elementes.

iii) Übung

iv) Offenbar ist wegen $\varphi(e) = e'$ also $e \in \mathrm{Ker}(\varphi)$. Weiter ist mit $x, y \in \mathrm{Ker}(\varphi)$ auch $xy^{-1} \in \mathrm{Ker}(\varphi)$, denn es ist $\varphi(xy^{-1}) = \varphi(x)\varphi(y)^{-1} = e$. Schließlich besagt die Normalteilereigenschaft, daß mit $x \in \mathrm{Ker}(\varphi)$ und $a \in G$ auch $axa^{-1} \in \mathrm{Ker}(\varphi)$ ist. Das ist richtig, denn es ist

$$\varphi(axa^{-1}) = \varphi(a)\varphi(x)\varphi(a)^{-1} = \varphi(a)e'\varphi(a)^{-1} = e'$$

□

13.13 Definition und Satz: Eine Untergruppe N einer Gruppe G ist ein Normalteiler, wenn eine der folgenden äquivalenten Bedingungen erfüllt ist:

i) Für alle $g \in G$ ist $gN = \{gn \mid n \in N\} = Ng = \{ng \mid n \in N\}$

ii) Es ist $gNg^{-1} := \{gng^{-1} \mid n \in N\} = N$

Beweis: Übung! □

13.14 Beispiele:

i) Sei $G = S_n$, $G' = \{\pm 1\}$ mit Multiplikation als Verknüpfung.
$\mathrm{sgn} : G \to G'$, $\pi \mapsto \mathrm{sgn}(\pi)$ ist ein Gruppenhomomorphismus. $\mathrm{Ker}(\mathrm{sgn}) = A_n \subset S_n$ ist die alternierende Gruppe.

ii) Sei V ein endlich dimensionaler Vektorraum über dem Körper K, $G = \mathrm{GL}(V)$ die allgemein lineare Gruppe der Automorphismen des Vektorraumes V. Dann ist die Determinantenabbildung $\det : \mathrm{GL}(V) \to (K^\times, \cdot)$ ein Gruppenhomomorphismus in die multiplikative Gruppe des Körpers K. Es ist

$$\mathrm{Ker}(\det) = \{\varphi \in \mathrm{GL}(V) \mid \det(\varphi) = 1\} =: \mathrm{SL}(V)$$

die sogenannte spezielle lineare Gruppe, ein Normalteiler in der $\mathrm{GL}(V)$.

iii) Sei $\varphi : V \to W$ ein Vektorraumhomomorphismus. Dann ist φ insbesondere auch ein Homomorphismus der Gruppe $(V, +)$ in die Gruppe $(W, +)$.

13.15 Satz: Sei G eine Gruppe, $a \in G$. Dann ist

$$\left.\begin{array}{rcccl} \mathrm{Int}(a) & : & G & \to & G \\ & & x & \mapsto & axa^{-1} \end{array}\right\}$$

ein sogenannter innerer Automorphismus der Gruppe G.

Beweis: Übung. □

13.16 Übung:

i) Sei $\mathrm{Aut}(G)$ die Menge der Automorphismen der Gruppe G. Diese bildet selber bezüglich Komposition eine Gruppe.

ii) Die Abbildung $\mathrm{Int} : G \to \mathrm{Aut}(G)$, $a \mapsto \mathrm{Int}(a)$ ist ein Homomorphismus von Gruppen.

13.17 Definition: Sei V ein endlich dimensionaler Vektorraum. Eine *Translation* von V mit $v \in V$ ist die Abbildung $T_v : V \to V, x \mapsto T_v(x) = x + v$.

13.18 Proposition: Mit Hintereinanderausführung als Komposition bilden die Translationen eines Vektorraumes V eine Gruppe, die sogenannte Gruppe der Translationen. Diese ist isomorph zum Vektorraum V, aufgefaßt als Gruppe $(V, +)$.

Beweis: Es ist $T_{v_1} \circ T_{v_2} = T_{v_1+v_2} = T_{v_2} \circ T_{v_1}$ Die Zuordnung $v \mapsto T_v$ liefert den gewünschten Isomorphismus. $\square$

13.19 Definition: Eine affine Abbildung des zu V gehörenden affinen Raumes in den zum Vektorraum W gehörenden affinen Raum W, $A : V \to W$, wird gegeben durch Hintereinanderausführung einer linearen Abbildung φ und einer Translation T_v, d. h. $A = (T_v \circ \varphi)$. Explizit gilt für $x \in V$: $A(x) = \varphi(x) + v$.

13.20 Satz:

i) Die Komposition von zwei affinen Abbildungen $A : V \to W$, $B : W \to U$ ist wieder affin.

ii) Eine affine Abbildung ist bijektiv genau, wenn der „lineare Anteil" φ (siehe oben) bijektiv ist. Dann ist die inverse Abbildung ebenfalls affin.

iii) Die Menge aller invertierbaren affinen Abbildungen $A : V \to V$ bildet mit Komposition eine Gruppe, die sogenannte affine Gruppe $\mathrm{Aff}(V)$. $\mathrm{Aff}(V)$ enthält als Untergruppe die Gruppen $\mathrm{GL}(V)$ und die Gruppe T der Translationen. T ist sogar ein Normalteiler in $\mathrm{Aff}(V)$.

Beweis:

i) $A(x) = \varphi(x) + v$, $B(x) = \psi(x) + w$. Dann ist

$$(A \circ B)(x) = A(B(x)) = A(\psi(x) + w) = \varphi(\psi(x) + w) + v = (\varphi \circ \psi)(x) + (\varphi(w) + v)$$

Das heißt $(A \circ B)$ ist gegeben durch Hintereinanderausführung der linearen Abbildung $(\varphi \circ \psi)$ und der Translation $T_{\varphi(w)+v}$.

ii) Ist φ invertierbar und $y = \varphi(x) + v$, so ergibt sich

$$x = \varphi^{-1}(y - v) = \varphi^{-1}(y) - \varphi^{-1}(v) = (T_{-\varphi^{-1}(v)} \circ \varphi^{-1})(y)$$

Also ist die inverse Abbildung affin.

iii) Übung.

$\square$

14 Eigenwerte und Vektoren

Ist eine lineare Abbildung

$$\varphi : V \to V$$

eines K-Vektorraumes V durch eine Matrix gegeben, etwa

$$A = \mathrm{Matrix}(\varphi : \{e_1, \dots, e_n\})$$

bezüglich einer Basis $\{e_1, \dots, e_n\}$, sodass

$$A = \begin{pmatrix} \alpha_{11} & & 0 \\ & \ddots & \\ 0 & & \alpha_{nn} \end{pmatrix}$$

in Diagonalgestalt vorliegt, so kann man mit ihr in dieser Form besonders gut rechnen. Zum Beispiel kann man unmittelbar die Potenzen

$$A^k = \begin{pmatrix} \alpha_{11}^k & & 0 \\ & \ddots & \\ 0 & & \alpha_{nn}^k \end{pmatrix}$$

für $k \in \mathbb{Z}$ beliebig ausrechnen oder etwa die Determinante

$$\det(A) = \prod_{i=1}^{n} \alpha_{ii} \quad .$$

Für die Basisvektoren $\{e_1, \dots, e_n\}$ gilt

$$\varphi(e_i) = \alpha_{ii} e_i \quad (i = 1, \dots, n) \quad ,$$

d. h. die Vektoren e_i werden in skalare Vielfache von sich selbst abgebildet oder anders gesagt:

$$\varphi(Ke_i) \subseteq Ke_i \quad (i = 1, \dots, n)$$

d. h. die durch e_i bestimmte Gerade Ke_i wird in sich selbst abgebildet.

Aus diesen Gründen stellen wir uns jetzt folgende Frage: Gegeben eine beliebige lineare Abbildung

$$\varphi : V \to V$$

Gesucht sind Vektoren $x \in V$ und Skalare $\lambda \in K$, sodass gilt:

$$\varphi(x) = \lambda x$$

14.1 Definition: Sei $\varphi : V \to V$ eine lineare Abbildung des Vektorraumes V. Ein Eigenvektor x zum Eigenwert λ ist ein Vektor $x \in V$, $x \neq 0$, für den gilt:

$$\varphi(x) = \lambda x$$

Wir untersuchen jetzt, wie man Eigenwerte und Eigenvektoren finden kann.

Sei also $x \in V$, $x \neq 0$ und $\varphi(x) = \lambda x$ mit $\lambda \in K$. Wir schreiben dies als $\varphi(x) = \lambda\,\mathrm{id}(x)$ mittels der identischen Abbildung $\mathrm{id} : V \to V$. Daraus folgt $(\varphi - \lambda\,\mathrm{id})(x) = 0$, d. h. die lineare Abbildung $(\varphi - \lambda\,\mathrm{id})$ ist nicht injektiv, da sie den Vektor x, $x \neq 0$, auf 0 abbildet. Daraus folgt für die Determinante

$$\det(\varphi - \lambda\,\mathrm{id}) = 0$$

als Bedingung an den Eigenwert λ. Es gilt aber auch die Umkehrung!

Sei für $\lambda \in K$

$$\det(\varphi - \lambda\,\mathrm{id}) = 0$$

Dies bedeutet nach den allgemeinen Resultaten über lineare Abbildungen, dass die lineare Abbildung $\varphi - \lambda\,\mathrm{id} : V \to V$ nicht injektiv ist. Das bedeutet aber, dass $\mathrm{Ker}(\varphi - \lambda\,\mathrm{id}) \neq 0$ ist, bzw. es existiert $x \in V$, $x \neq 0$ mit $(\varphi - \lambda\,\mathrm{id})(x) = 0$

Letzteres ist aber äquivalent zu $\varphi(x) = \lambda x$, also ist $x \in V$ darum auch ein Eigenvektor zum Eigenwert λ.

Die Bedingung

$$\det(\varphi - \lambda\,\mathrm{id}) \neq 0$$

wollen wir jetzt noch etwas genauer verstehen.

Sei $\{e_1, \ldots, e_n\}$ eine beliebige Basis von V,

$$A := \mathrm{Matrix}(\varphi; \{e_1, \ldots, e_n\}) = (\alpha_{ij})_{1 \leq i,j \leq n} \quad ,$$

die zugehörige Matrix, die ganz beliebig sein kann.
Wir haben

$$\det(\varphi - \lambda\,\mathrm{id}) = \det(A - \lambda E) \quad ,$$

da $E = \mathrm{Matrix}(\mathrm{id}; \{e_1, \ldots, e_n\})$.
Es ist

$$(A - \lambda E) = (\alpha_{ij} - \lambda\delta_{ij})_{1 \leq i,j \leq n}$$

mit Kronecker-Symbol δ_{ij}.

Aus der Entwicklungsformel für die Determinante ergibt sich sofort

$$\det(A - \lambda E) = (-1)^n \lambda^n + (-1)^{n-1} \left(\sum_{i=1}^{n} \alpha_{ii} \right) \lambda^{n-1} + \cdots + \det(A) \quad .$$

D. h. , es ergibt sich ein Polynom, das sogenannte charakteristische Polynom in λ, mit höchstem Koeffizienten $(-1)^n$ und konstantem Koeffizienten $\det(A)$. Wir können auch schreiben

$$\det(\varphi - \lambda\,\mathrm{id}) = (-1)^n \lambda^n + (-1)^{n-1}\,\mathrm{Sp}(\varphi)\lambda^{n-1} + \cdots + \det(\varphi) \quad ,$$

das charakteristische Polynom hängt nach seiner Definition nur vom Endomorphismus φ ab.

Wir fassen die Ergebnisse zusammen.

14.2 Satz: Sei $\varphi : V \to V$ eine lineare Abbildung des endlichdimensionalen Vektorraums V. Ein Skalar $\lambda \in K$ ist Eigenwert von φ genau dann, wenn gilt: $\det(\varphi - \lambda\,\mathrm{id}) = 0$, d. h. genau, wenn λ eine Nullstelle des charakteristischen Polynoms

$$\det(\varphi - \lambda\,\mathrm{id}) = (-1)^n \lambda^n + (-1)^{n-1}\,\mathrm{Sp}(\varphi)\lambda^{n-1} + \cdots + \det(\varphi)$$

ist. Zu einem Eigenwert $\lambda \in K$ gibt es dann insbesondere einen Eigenvektor $x \in V$, $x \neq 0$, der die Eigenwertgleichung

$$\varphi(x) = \lambda x$$

erfüllt.

14.3 Bemerkungen: Man kann die obigen Betrachtungen natürlich auch direkt für Matrizen durchführen. Gesucht ist dann zu einer $(n \times n)$-Matrix $A \in \mathbb{M}(n, K)$, $A = (\alpha_{ij})_{1 \leq i,j \leq n}$ ein Vektor $x = (\xi_1, \ldots, \xi_n)^t$, sodass gilt

$$\begin{aligned} A \cdot x &= \lambda x \\ &= \lambda E \cdot x \end{aligned}$$

mit $E \in \mathbb{M}(n; K)$ die Einheitsmatrix. Daraus ergibt sich wieder

$$\det(A - \lambda E) = 0 \quad ,$$

d. h. λ muss Nullstelle des charakteristischen Polynoms sein.

14.4 Beispiele:

i) Sei

$$A = \begin{pmatrix} \alpha_1 & & 0 \\ & \ddots & \\ 0 & & \alpha_n \end{pmatrix}$$

eine Diagonalmatrix mit Diagonaltermen $\alpha_1, \ldots, \alpha_n \in K$. Dann sind $\alpha_1, \ldots, \alpha_n$ auch genau die Eigenwerte der zu A gehörenden linearen Abbildung.
$e_i \quad (i = 1, \ldots, n)$ ist jeweils Eigenvektor zum Eigenwert λ_i, allerdings müssen dies keineswegs alle Eigenvektoren sein, siehe hierzu Beispiel (ii).

ii) Sei $A = \alpha E$, $\alpha \in K$.
Dann ist *jeder* Vektor $x \in K^n$ ein Eigenvektor zum Eigenwert $\lambda = \alpha$. Entsprechendes gilt natürlich für einen Endomorphismus $\varphi = \alpha \cdot \mathrm{id}$.

iii) Eine lineare Abbildung hat eventuell keinen Eigenwert. Sei etwa

$$A = \begin{pmatrix} \cos(\varphi) & \sin(\varphi) \\ -\sin(\varphi) & \cos(\varphi) \end{pmatrix} : \mathbb{R}^2 \to \mathbb{R}^2$$

eine reelle Matrix, die die Drehung um einen Winkel φ beschreibt. Ist etwa $0 < \varphi < \pi$, so wird offenbar keine Gerade durch den Nullpunkt bei der Drehung in sich übergeführt. In der Tat ergibt sich für das charakteristische Polynom

$$\lambda^2 - 2\lambda\cos(\varphi) + 1 = 0 \quad \text{bzw.}$$

$$(\lambda - \cos(\varphi))^2 = -1 + \cos^2(\varphi) < 0 \quad \text{für } 0 < \varphi < \pi$$

Daher hat das charakteristische Polynom keine reelle Lösung, was den oben genannten anschaulichen Befund bestätigt.

Kurzes Intermezzo: Polynome und Polynomfunktionen: K sei Körper; eine *Polynomfunktion* ist eine Abbildung

$$p : K \to K, \quad \lambda \mapsto p(\lambda) = \sum_{i=0}^{n} \alpha_i \lambda^i$$

dabei $\{\alpha_0, \ldots, \alpha_n\} \subset K$ fest.
Ein *Polynom* ist ein formaler Ausdruck der Form $\sum_{i=0}^{n} \alpha_i x^i$.

Die Menge aller Polynome mit Koeffizienten $\alpha_i \in K \quad (i = 0, \ldots, n)$ bildet mit der üblichen formalen Addition und Multiplikation einen Ring, den sogenannten *Polynomring* $K[x]$ über K. Offenbar bestimmt jedes Polynom $p \in K[x]$, $p = \sum_{i=0}^{n} \alpha_i x^i$ eine zugehörige Polynomfunktion

$$\overline{p} = p : K \to K, \quad \lambda \mapsto \sum_{i=0}^{n} \alpha_i \lambda^i \quad .$$

Umgekehrt bestimmt $\overline{p}$ eindeutig das Polynom p, sofern K unendlich ist. Ist K dagegen ein endlicher Körper, so gibt es offenbar nur $|K|^{|K|}$ verschiedene Funktionen $K \to K$, also auch höchstens so viele Polynomfunktionen, dagegen gibt es unendlich viele Polynome in $K[x]$.

Beispiel: Ist $K = \mathbb{F}_2$, so ist die Polynomfunktion $p : \mathbb{F} \to \mathbb{F}, \quad \lambda \mapsto (\lambda^2 - \lambda)$ offenbar identisch 0, also $p = 0$ als Polynomfunktion. Dagegen sind die Polynome 0 und $(x^2 - x)$ in $K[x]$ verschiedene Elemente.

14.5a Definition: Ein Polynom $f(x) \in K[x]$ (dem Polynomring über K) heißt irreduzibel, wenn es keine *echte* Zerlegung $f(x) = f_1(x)f_2(x)$ (mit $\deg(f_1) < \deg(f)$, $\deg(f_2) < \deg(f)$) gibt.

14.5b Satz: Jedes Polynom $f(x)$ kann als Produkt von irreduziblen Polynomen geschrieben werden. Diese Zerlegung ist im Wesentlichen (d. h. bis auf Multiplikation mit (konstanten) Elementen aus K^* und bis auf die Anordnung der Faktoren) eindeutig. Diesen Satz werden wir erst im zweiten Teil der Vorlesung beweisen. Das Ergebnis werden wir allerdings unten benutzen.

Für die Nullstellen einer Polynomfunktion gelten folgende Sätze.

14.5c Satz: Sei $P(x) = \sum_{i=0}^{n} p_i x^i$, $p_i \in K$ für $i = 0, \ldots, n$ ein Polynom über K, $\alpha \in K$ sei Nullstelle von P, d. h. $P(\alpha) = \sum_{i=0}^{n} p_i \alpha^i = 0$. Dann kann man das Polynom $P(x)$ schreiben als

$$P(x) = (x - \alpha) P_1(x)$$

Beweis: Es ist

$$\begin{aligned} P(x) &= P(x) - P(\alpha) \\ &= \sum_{i=0}^{n} p_i (x^i - \alpha^i) \\ &= \left(\sum_{i=0}^{n} p_i (x^{i-1} + \cdots + \alpha^{i-1}) \right) (x - \alpha) \end{aligned}$$

$\square$

14.6 Bemerkung: Sei $P(x) = \sum_{i=0}^{n} p_i x^i$ wie oben, $\alpha \in K$. α ist eine Nulllstelle m-ter Ordnung von f $:\Leftrightarrow$ Man kann schreiben $P(x) = (x - \alpha)^m P_1(x)$ mit $P_1(\alpha) \neq 0$. m heißt auch die *Vielfachheit* oder *Multiplizität* der Nullstelle α.

14.7 Satz: Sei $P(x) = \sum_{i=0}^{n} p_i x^i$ wie oben ein beliebiges Polynom vom Grad n, $\alpha_1, \ldots, \alpha_r$ seien die sämtlichen Nullstellen von $P(x)$, $m_1, \ldots, m_r$ ihre Multiplizitäten. Dann gilt: Es ist $\sum_{i=1}^{r} m_i \leq n = \deg P(x)$, d. h. es gibt für jedes Polynom n-ten Grades höchstens n Nullstellen, wobei Vielfachheiten berücksichtigt sind. Speziell gibt es höchstens n verschiedene Nullstellen von P in K.

Beweis: Schreibe $P(x)$ als Produkt seiner irreduziblen Faktoren. Aufgrund der Eindeutigkeit treten auf jeden Fall die Faktoren $(x - \alpha_i)^{m_i}$ alle auf. Daher hat man eine Zerlegung

$$P(x) = \prod_{i=1}^{r} (x - \alpha_i)^{m_i} P_1(x)$$

Die Behauptung folgt. $\square$

14.8 Definition: Ein Körper K heißt *algebraisch abgeschlossen* $:\Leftrightarrow$ Jedes Polynom vom Grad n,

$$P(x) = \sum_{i=0}^{n} p_i x^i \quad (\text{mit } p_n = c, c \neq 0)$$

kann geschrieben werden als Produkt von Linearfaktoren mit Vielfachheiten, also

$$P(x) = c \cdot \prod_{i=1}^{r} (x - \alpha_i)^{m_i}, \quad c \in K^x$$

Speziell gilt dann $\sum_{i=1}^{r} m_i = n$.

14.9 Satz: Der Körper $\mathbb{C}$ der komplexen Zahlen ist algebraisch abgeschlossen.

Beweis: Siehe Vorlesung zur Algebra ($\rightarrow$ algebraischer Beweis) bzw. Funktionentheorievorlesung ($\rightarrow$ analytischer Beweis).

Bemerkungen:

i) Tatsächlich gibt es zu jedem Körper K einen Erweiterungskörper $\tilde{K}$, der K enthält und algebraisch abgeschlossen ist. Der bis auf Isomorphie eindeutig bestimmte kleinste derartige Körper ist der (ein) *algebraischer Abschluss* $\overline{K}$ von K.

ii) Zum Beispiel ist der algebraische Abschluss des Körpers $\mathbb{R}$ der reellen Zahlen der Körper $\mathbb{C}$ der komplexen Zahlen.

iii) Über $\mathbb{R}$ kann man immerhin jedes Polynom

$$P(x) = P_1(x) \cdot \dots \cdot P_r(x)$$

noch als Produkt von linearen und quadratischen Polynomen schreiben.

14.10 Satz: Hat der Endomorphismus $\varphi : V \rightarrow V$ des n-dimensionalen K-Vektorraums V n verschiedene Eigenwerte $\lambda_1, \dots, \lambda_n$, so hat φ bezüglich einer Basis aus Eigenvektoren $\{e_1, \dots, e_n\}$ die Gestalt

$$\mathrm{Matrix}(\varphi; \{e_1, \dots, e_n\}) = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix} \quad ,$$

hat also Diagonalgestalt.

Beweis: Zu jedem der $\lambda_i \quad (i = 1, 2, \dots, n)$ findet man $e_i \in V$, $e_i \neq 0$ mit $\varphi(e_i) = \lambda_i e_i$. Es bleibt zu zeigen: $\{e_1, \dots, e_n\}$ ist ein linear unabhängiges System. Sei etwa (O.E. durch Umnumerieren) $\{e_1, \dots, e_r\}$ mit r minimal, so dass eine nichttriviale Relation

$$\mu_1 e_1 + \mu_2 e_2 + \dots + \mu_r e_r = 0$$

besteht. Speziell sind alle $\mu_i \neq 0$.

Anwendung des Endomorphismus φ liefert

$$\mu_1 \lambda_1 e_1 + \mu_2 \lambda_2 e_2 + \dots + \mu_r \lambda_r e_r = 0$$

Dabei sind die λ_i alle voneinander verschieden für $i = 1, 2, \ldots, r$. Wir haben weiter

$$\lambda_1\mu_1 e_1 + \ldots + \lambda_1\mu_r e_r = 0$$

Substraktion ergibt eine Relation

$$(\mu_2\lambda_2 - \mu_2\lambda_1)e_2 + \ldots + (\mu_r\lambda_r - \mu_r\lambda_1)e_r = 0$$

Da r minimal war, folgt:

$$\mu_2(\lambda_2 - \lambda_1) = \ldots = \mu_r(\lambda_r - \lambda_1) = 0$$

Da aber die λ_i paarweise verschieden sind, folgt:

$$\mu_2 = \ldots = \mu_r = 0$$

Dann folgt aber $\mu_1 e_1 = 0$, wegen $e_1 \neq 0$ also auch $\mu_1 = 0$. Also war die Ausgangsrelation entgegen der Annahme trivial. Widerspruch. $\square$

14.11 Definition und Satz: Sei $\varphi : V \to V$ ein Endomorphismus, $\lambda \in K$. Der *Eigenraum* zum *Eigenwert* λ ist

$$V(\lambda; \varphi) := \{x \in V | \varphi(x) = \lambda x\}$$

Insbesondere ist $V(\lambda; \varphi)$ ein Untervektorraum von V.

Beweis: Übung. $\square$

14.12 Definition und Satz: Seien φ, λ wie oben. Der verallgemeinerte Eigenraum zu λ ist

$$\tilde{V}(\varphi; \lambda) := \{x \in V | \exists N > 0, (\varphi - \lambda\,\mathrm{id})^N(x) = 0\}$$

$\tilde{V}(\varphi; \lambda)$ ist wiederum ein Untervektorraum von V.

Beweis: Seien $x, y \in \tilde{V}(\varphi; \lambda)$, etwa $(\varphi - \lambda\,\mathrm{id})^N(x) = 0$, $(\varphi - \lambda\,\mathrm{id})^M(y) = 0$ mit $N, M \in \mathbb{N}$ geeignet. Sei o.E.d.A. $N = max\{N, M\}$. Dann ist auch

$$(\varphi - \lambda\,\mathrm{id})^N(y) = 0\ ,$$

damit aber auch $(\varphi - \lambda\,\mathrm{id})^N(x + y) = 0$. Entsprechend natürlich $(\varphi - \lambda\,\mathrm{id})^N(\alpha x) = 0$. $\square$

14.12' Satz: Der Operator φ (beziehungsweise die ganze Algebra $\mathbb{C}[\varphi]$ der Endomorphismen $\sum_{j=0}^m \lambda_j \varphi^j$) bildet die Eigenräume und die verallgemeinerten Eigenräume in sich ab.

Beweis: Wir zeigen zum Beispiel

$$\varphi(\tilde{V}(\lambda; \varphi)) \subseteq \tilde{V}(\lambda; \varphi)$$

Sei also $v \in \tilde{V}(\lambda;\varphi)$. Dann existiert $N > 0$, $(\varphi-\lambda)^N(v) = 0$. Dann ist aber

$$\begin{aligned} & (\varphi-\lambda)^N\,(\varphi(v)) \\ = \; & \varphi(\varphi-\lambda)^N(v) \quad \text{(da } \mathbb{C}[\varphi] \text{ kommutativ ist)} \\ = \; & \varphi(0) = 0 \end{aligned}$$

Also folgt: $\varphi(v) \in \tilde{V}(\lambda;\varphi)$. $\square$

14.13 Satz: Seien $\lambda_1, \ldots, \lambda_r \in K$, paarweise verschieden. Dann ist

i)

$$\sum_{i=1}^{r} V(\varphi;\lambda_i) = \bigoplus_{i=1}^{r} V(\varphi;\lambda_i)$$

eine direkte Zerlegung.

ii)

$$\sum_{i=1}^{r} \tilde{V}(\varphi;\lambda_i) = \bigoplus_{i=1}^{r} \tilde{V}(\varphi;\lambda_i)$$

ebenfalls eine direkte Zerlegung.

Beweis: Wir zeigen, wie die Aussage (ii) auf (i) reduziert wird.

Angenommen, es besteht eine nichttriviale Relation der Form $\sum_{i=1}^{r'} v_i = 0$, dabei $v_i \in \tilde{V}(\varphi;\lambda_i) : r'$ sei minimal. Wenden wir etwa $(\varphi - \lambda_j)$ an, so ergibt sich

$$\sum_{i=1}^{r'} (\varphi-\lambda_j)(v_i) = 0 \quad (j \in \{1,\ldots,r\})$$

Da nach **14.12'** $(\varphi-\lambda_j)(v_i) \in \tilde{V}(\varphi;\lambda_i)$ für $i = 1,\ldots,r$ gilt, können wir durch sukzessive Anwendung der Operatoren $(\varphi - \lambda_j\,\mathrm{id})$ erreichen, dass generell $(\varphi-\lambda_i)(v_i) = 0$ gilt. D. h. aber, wir haben (ii) auf (i) reduziert. Aber (i) wird genau wie **14.10** gezeigt. $\square$

14.14 Ein typisches Beispiel: Sei $\lambda \in K$, wir betrachten die Matrix

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

mit λ's in der Diagonalen, 1 in der Nebendiagonalen. Wir berechnen $V(\lambda;\varphi)$, $\tilde{V}(\lambda;\varphi)$ und zeigen $V(\beta;\varphi) = 0$ falls $\beta \neq \lambda$:

a)
$$V(\lambda;\varphi) = K \cdot e_1$$

Beweis: Nachrechnen!

b) Offenbar ist $\tilde{V}(\lambda;\varphi) = V$, da ja
$$(A - \lambda\,\mathrm{id}) = \begin{pmatrix} 0 & 1 & & 0 \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ 0 & & & 0 \end{pmatrix}$$
und weiter gilt:
$$\begin{pmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ & & & \ddots & \\ 0 & & & 0 & 1 \\ 0 & & & 0 & 0 \end{pmatrix}^n = 0$$

c) Offenbar ist $V(\beta;\varphi) = 0$. Dann ist auch $\tilde{V}(\beta;\varphi) = 0$.

14.15 Satz: (*Jordansche Normalform*) Sei $\varphi : V \to V$ ein Endomorphismus des Vektorraums V über dem algebraisch abgeschlossenen Körper K. Dann besitzt V eine direkte Zerlegung $V = \bigoplus_{i=1}^r V_i$, invariant unter φ, so dass für jedes V_i eine Basis $\left\{e_1^{(i)}, \dots, e_{n_i}^{(i)}\right\}$ existiert mit
$$\left.\begin{array}{rcl} \varphi\left(e_1^{(i)}\right) & = & \lambda_i e_1^{(i)} \\ \varphi\left(e_2^{(i)}\right) & = & \lambda_i e_2^{(i)} + e_1^{(i)} \\ \vdots & & \vdots \\ \varphi\left(e_{n_i}^{(i)}\right) & = & \lambda_i e_{n_i}^{(i)} + e_{n_i-1}^{(i)} \end{array}\right\}$$
Diese Zerlegung ist im Wesentlichen eindeutig.

Beweis: (Durch vollständige Induktion nach der Dimension $\dim(V)$ von V)
Der Fall $\dim(V) = 1$: Hier ist alles trivial.

Induktionsschritt: Sei also $\varphi : V \to V$ gegeben, $\lambda \in K$ sei Eigenwert. Wir betrachten $\varphi - \lambda\,\mathrm{id} : V \to V$. Wegen $\mathrm{Ker}(\varphi - \lambda\,\mathrm{id}) \neq 0$ folgt:
$$\tilde{V} := (\varphi - \lambda\,\mathrm{id})(V) \subsetneq V \quad \text{Weiter ist}$$
$$\varphi(\tilde{V}) \subseteq \tilde{V}\ .$$
Also können wir nach Induktionsannahme $\varphi|\tilde{V}$ in Jordanscher Normalform annehmen. Das heißt, wir finden $\tilde{V}_i$, $i = 1, \dots, \tilde{r}$ mit:

1) Es ist $\tilde{\varphi}(\tilde{V}_i) \subseteq \tilde{V}_i$, d. h. die $\tilde{V}_i$ sind Eigenräume von φ.

2)
$$\text{Es ist} \quad \tilde{V} = \bigoplus_{i=1}^{\tilde{r}} \tilde{V}_i$$

3) Es gibt Basis $\left\{e_1^{(i)}, \ldots, e_{\tilde{r}_i}^{(i)}\right\}$ von $\tilde{V}_i$ mit

$$\begin{cases} \varphi\left(e_1^{(i)}\right) & = & \lambda_i e_1^{(i)} \\ \varphi\left(e_2^{(i)}\right) & = & \lambda_i e_2^{(i)} + e_1^{(i)} \\ \vdots & & \vdots \\ \varphi\left(e_{\tilde{r}_i}^{(i)}\right) & = & \lambda_i e_{\tilde{r}_i}^{(i)} + e_{\tilde{r}_i-1}^{(i)} \end{cases}$$

Ist $\lambda_i \neq \lambda$ so sieht man sofort: $\varphi - \lambda \operatorname{id}|_{V_i}$ ist bijektiv (der einzige Eigenwert von $(\varphi|_{V_i})$ ist ja doch $\lambda_i \neq \lambda$).
Ist $\lambda = \lambda_i$ für $i \in \{1, \ldots, \tilde{s}\}$, etwa für $i \in \{1, \ldots, t\}$ mit $t \leq \tilde{s}$, so wähle jeweils $e_{\tilde{r}_i+1}^{(i)}$ mit

$$(\varphi - \lambda_i)\left(e_{\tilde{r}_i+1}^{(i)}\right) = e_{\tilde{r}_i}^{(i)}$$

Wir setzen

$$\tilde{r}_{i+1} =: r_i \text{ für diese } i \text{ mit } \lambda_i = \lambda, \quad \tilde{r}_i =: r_i \quad \text{für die anderen.}$$

Wir ergänzen weiter das System der Vektoren $\left\{e_1^{(1)}, e_1^{(2)}, \ldots, e_1^{(t)}\right\}$, das ja linear unabhängig ist nach Induktionsvorraussetzung und aus $\operatorname{Ker}(\varphi - \lambda \operatorname{id})$ ist, zu einer Basis von $\operatorname{Ker}(\varphi - \lambda \operatorname{id})$. Die neuen Vektoren seien etwa

$$\left\{e_1^{(t+1)}, e_1^{t+2)}, \ldots, e_1^{(t+r)}\right\}.$$

Behauptung: Das System der Vektoren

$$\left\{e_i^{(j)} | i \in \{1, \ldots, \tilde{s}\},\, j \in \{1, \ldots, r_i\}\right\} \cup \left\{e_1^{(t+1)}, e_1^{(t+2)}, \ldots, e_1^{(t+r)}\right\}$$

(mit $s := t + r$)
liefert eine Basis von V, bezüglich der φ in Jordanscher Normalform ist.

Beweis: Zunächst zeigen wir: Obiges System ist ein Erzeugendensystem von V. Dies ist klar, denn wir schreiben für einen beliebigen Vektor $v \in V$

$$(\varphi - \lambda)(v) = \sum_{i=1}^{\tilde{s}} \sum_{h=1}^{\tilde{r}_i} \lambda_{ih} e_h^{(i)}$$

Da aber jeder der Vektoren $e_h^{(i)}$ nach obiger Überlegung in der Form

$$(\varphi - \lambda)(x_i) \quad \text{mit} \quad x_i \in \sum_{k=1}^{r_i} K e_k^{(i)}$$

geschrieben werden kann, folgt: Wir finden eine Linearkombination der oben angegebenen Basisvektoren,

$$y := \sum_{i=1}^{\tilde{s}} \sum_{j=1}^{r_i} \mu_{ij} e_i^{(j)}$$

mit

$$(\varphi - \lambda)(v) = (\varphi - \lambda)(y)$$

D. h. aber: $(\varphi - \lambda)(v - y) = 0$.
Dann kann aber $(v - y)$ als Linearkombination der Vektoren aus

$$\left\{ e_1^{(1)}, \ldots, e_1^{(s)} \right\}$$

geschrieben werden.
Es folgt, dass wir oben ein Erzeugendensystem vorliegen haben.

Als nächsten Schritt zeigen wir, dass unser System oben *linear unabhängig* ist. Unser obiges System besteht aus

$$\sum_{i=1}^{\tilde{s}} r_i + r \qquad \text{Vektoren.}$$

Es ist

$$\dim(\varphi - \lambda \, \mathrm{id})(V) = \sum_{i=1}^{t} (r_i - 1) + \sum_{i=t+1}^{\tilde{s}} r_i$$

weiter ist

$$\dim \, \mathrm{Ker}(\varphi - \lambda) = (t + r).$$

Damit folgt aber:

$$\dim(V) = \sum_{i=1}^{t} r_i + \sum_{i=t+1}^{\tilde{s}} r_i + r \ ,$$

also genau die Anzahl der Vektoren unseres oben angegebenen Systems. Daher ist unser System nicht nur ein Erzeugendensystem, sondern auch notwendig linear unabhängig, also eine Basis. Damit ist die Existenz einer Jordanschen Normalform gezeigt.

Eindeutigkeit: Die Menge der auftretenden Eigenwerte $\{\lambda_i\}$ ist durch φ natürlich eindeutig bestimmt. Die Dimensionen

$$\dim \left(\mathrm{Ker}(\varphi - \lambda_i \, \mathrm{id})^j \right), \quad i = 1, \ldots, s \ , \quad j = 1, 2, \ldots$$

bestimmen, wie man leicht nachrechnet, die sämtlichen numerischen Paramter der Jordanschen Zerlegung eindeutig. Die Behauptung folgt. □

Wir geben zum besseren Verständnis noch die entsprechende Aussage über Matrizen, die natürlich ohne weiteres aus obigem Resultat folgt.

Sei φ bezüglich einer Basis $\{e'_1, \ldots, e'_n\}$ durch eine Matrix A' gegeben, so wird φ bezüglich der passenden Basis $\{e_1, \ldots, e_n\}$ durch die Matrix

$$XA'X^{-1} = A$$

gegeben.

Dabei beschreibt X den Übergang von der Basis $\{e'_1, \ldots, e'_n\}$ zur Basis $\{e_1, \ldots, e_n\}$. Die Matrizen A', A heißen zueinander konjugiert. „Konjugiertheit“ ist eine Äquivalenzrelation, die zueinander konjugierten Elemente bilden sogenannte Konjugationsklassen. Wir können daher auch so formulieren:

14.16 Satz: Jede $(n \times n)$-Matrix $A' \in \mathbb{M}(n;k)$ über einem algebraisch abgeschlossenen Körper K ist konjugiert zu einer Matrix A, die sich aus diagonalen Blöcken A_i zusammensetzt,

$$A = \begin{pmatrix} A_1 & & \\ & \ddots & \\ & & A_r \end{pmatrix}$$

dabei sind die A_i Jordanblöcke der Form

$$A_i = \begin{pmatrix} \lambda_i & 1 & & 0 \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ 0 & & & \lambda_i \end{pmatrix}$$

mit λ_i in der Diagonalen und 1 in der Nebendiagonalen. Diese Zerlegung ist im Wesentlichen eindeutig.

14.17 Anwendungen: Prozesse in der Physik führen oft auf Systeme lineare Differentialgleichungen

$$\left.\begin{array}{rcl} \dfrac{d\xi_1}{dt} & = & \displaystyle\sum_{j=1}^{n} \alpha_{1j}(\xi_1, \ldots, \xi_n)\xi_j \\ \vdots & & \vdots \\ \dfrac{d\xi_n}{dt} & = & \displaystyle\sum_{j=1}^{n} \alpha_{nj}(\xi_1, \ldots, \xi_n)\xi_j \end{array}\right\}$$

gesucht sind Lösungen

$$x = x(t) = (\xi_1(t), \ldots, \xi_n(t))$$

im $\mathbb{R}^n$ (eventuell als Spalten) geschrieben. In kompakter Form also:

$$\frac{d\xi_i}{dt} = \sum_{j=1}^{n} \alpha_{ij}(\xi_1, \ldots, \xi_n)\xi_j \quad (i = 1, \ldots, n)$$

oder in Matrixform

$$\frac{dx}{dt} = A \cdot x \quad \text{mit} \quad A = (\alpha_{ij}(x))$$

Man kann häufig A entwickeln nach den $\xi_1, \ldots, \xi_n$ oder zusammengefasst nach x. Etwa

$$A = A(x) = A_0 + A_1(x) + A_2(x) + \ldots$$

Dabei enthählt A_0 die konstanten Terme in $\xi_1, \ldots, \xi_n$, $A_1(x)$ die Terme erster Ordnung in $\xi_1, \ldots, \xi_n$ u. s. w. . In 0-ter Näherung hat man daher das System von Differentialgleichungen

$$\frac{dx}{dt} = Ax$$

mit $A = A_0$ eine konstante Matrix zu betrachten. Zur Anfangsbedingung $x = x(t_0) = x(0) \quad (t_0 = 0)$ ergibt sich eine Lösung der Form

$$\begin{aligned} x(t) &= \exp(t \cdot A)(x((0)) \quad \text{mit} \\ x(0) &= \left(\xi_1^{(0)}, \ldots, \xi_n^{(0)}\right)^t \end{aligned}$$

Es geht um das Verhalten von $\exp(tA)$ zum Beispiel für $t \to \infty$.

Bringe A auf Jordansche Normalform

$$\begin{aligned} A &= (C^{-1}\tilde{A}C), \text{ dann} \\ x(t) &= C^{-1}\exp(t\tilde{A})(C(x(0))) \\ \Rightarrow Cx(t) &= \exp(t\tilde{A})C(x(0)) \end{aligned}$$

Setze

$$C \cdot x(t) =: y(t), \quad \text{dann folgt}$$

$$y(t) = \exp(t\tilde{A}) \underbrace{(y(0))}_{\displaystyle \| \atop \displaystyle (\xi_0^{(0)}, \ldots, \xi_n^{(0)})}$$

15 Reelle Vektorräume und euklidische Geometrie

Bis hierher haben in der Vorlesung die Begriffe der (euklidischen) Geometrie wie Abstand und Winkel keine Rolle gespielt. Im letzten Teil der Vorlesung wollen wir daher gerade auf diese eingehen. Dafür ist es nötig, von nun an über dem Körper der reellem Zahlen bzw. mit geringen Modifikationen über dem Körper der komplexen Zahlen zu arbeiten. Tatsächlich kann man vieles dann auch wieder über beliebigen Körpern behandeln, wenn man mit allgemeinen Bilinearformen arbeiten will. Wir wollen aber hier, um den Zusammenhang zur Geometrie leichter zugänglich zu machen, wie oben beschrieben vom Speziellen zum Allgemeinen übergehen.

Wir werden dabei feststellen, daß die übliche euklidische Geometrie ohne Abstriche auch für höher dimensionale reelle Vektorräume ihre Richtigkeit behält.

V sei ein (in der Regel bei uns zunächst endlich dimensionaler) reeller Vektorraum.

15.1 Definition: Ein Skalarprodukt $(\quad,\quad)$ auf V ist eine Abbildung

$$\begin{aligned}(\quad,\quad) : V \times V &\to \mathbb{R}\\ (x,y) &\mapsto (x,y)\end{aligned}$$

mit folgenden Eigenschaften:

i) $(\quad,\quad)$ ist bilinear in den Variablen x, y. Das heißt also wie üblich, es ist

$$\begin{aligned}(\alpha_1 x_1 + \alpha_2 x_2, y) &= \alpha_1(x_1, y) + \alpha_2(x_2, y)\\ (x_1, \beta_1 y_1 + \beta_2 y_2) &= \beta_1(x, y_1) + \beta_2(x, y_2)\end{aligned}$$

für alle $\alpha_1, \alpha_2, \beta_1, \beta_2 \in \mathbb{R},\ x_1, x_2, y_1, y_2 \in V$.

ii) $(\quad,\quad)$ ist symmetrisch, d. h. es gilt $(x, y) = (y, x)$ für alle $x, y \in V$.

iii) (*Positive Definitheit*) Es ist $(x, x) \geq 0$ für alle $x \in V$ mit Gleichheit genau, wenn $x = 0$ ist.

15.2 Beispiele:

i) $V = \mathbb{R}^n$, $x = (\xi_1, \ldots, \xi_n)$, $y = (\eta_1, \ldots, \eta_n)$, dann ist

$$(x, y) := \sum_{i=1}^{n} \xi_i \eta_i$$

das übliche Skalarprodukt im n-dimensionalen Raum.

ii) Ein unendlich dimensionales Beispiel: $V := C(I)$ sei der Raum der stetigen reellwertigen Funktionen auf dem kompakten Intervall $I = [a, b]$. Auf V hat man das Skalarprodukt

$$(f, g) := \int_a^b f(x)g(x)dx$$

Dies wird erlauben, auf Räumen von Funktionen die geometrische Sprechweise anzuwenden.

15.3 Definition: Ein endlich dimensionaler euklidischer Vektorraum ist ein endlich dimensionaler reeller Vektorraum V, versehen mit einem Skalarprodukt $(\quad,\quad)$, Schreibweise $(V, (\quad,\quad))$

15.4 Definition: Die Länge eines Vektors $x \in V$ in einem endlich dimensionalen euklidischen Vektorraum V ist gegeben durch

$$\|x\| := (x, x)^{\frac{1}{2}} = \sqrt{(x, x)}$$

15.5 Satz: (*Cauchy-Schwarzsche Ungleichung*) Es ist für $x, y \in V$ beliebig

$$|(x, y)| \leq \|x\|\|y\|$$

mit Gleichheit genau, wenn x und y linear abhängig sind.

Beweis: Sei o.B.d.A. $y \neq 0$. Wir betrachten den Ausdruck ($t \in \mathbb{R}$)

$$(x + ty, x + ty) = p(t)$$

Offenbar ist

$$(x + ty, x + ty) \geq 0$$

Andererseits ist

$$p(t) = t^2(y, y) + 2t(x, y) + (x, x)$$

eine quadratische Funktion in t, die ihr Maximum für $t = -\frac{(x,y)}{(y,y)}$ annimmt. Dies eingesetzt, ergibt sich

$$\frac{(x, y)^2}{(y, y)^2}(y, y) - \frac{2(x, y)^2}{(y, y)} + (x, x) \geq 0$$

d. h.

$$(x, y)^2 \leq (x, x)(y, y)$$

mit Gleichheit nur, wenn

$$x - \frac{(x, y)}{(y, y)}y = 0$$

ist. Sind umgekehrt x, y linear abhängig, so folgt sofort Gleichheit in der Cauchy-Schwarzschen Ungleichung. $\square$

15.6 Satz: Die Norm (Länge) erfüllt folgende Eigenschaften

i) $\|x\| \geq 0$ mit Gleichheit genau, wenn $x = 0$ ist

ii) $\|\lambda x\| =| \lambda | \, \|x\|$

iii) $\|x+y\| \leq \|x\| + \|y\|$ mit Gleichheit nur dann, wenn x und y linear abhängig sind.

Beweis: Nur (iii) ist nicht sofort klar. Aber wir haben

$$(\|x\| + \|y\|)^2 = (x,x) + 2(x,x)^{\frac{1}{2}}(y,y)^{\frac{1}{2}} + (y,y) \geq (x+y, x+y) = (x,x) + (y,y) + 2(x,y)$$

wegen $(x,x)^{\frac{1}{2}}(y,y)^{\frac{1}{2}} \geq | (x,y) |$ mit Gleichheit genau bei linearer Abhängigkeit. □

15.7 Definition: Ein reeller Vektorraum V versehen mit einer Norm $\| \star \|$, die die obigen Eigenschaften (i),(ii),(iii) (ohne die Zusatzeigenschaft über genaue Gleichheit in (iii)) erfüllt, heißt normierter Vektorraum.

15.8 Bemerkung: Es gibt allerdings „viel mehr“ Normen auf einem Vektorraum als Skalarprodukte, d. h. die wenigsten Normen auf einem Vektorraum V kommen von einem Skalarprodukt her (siehe Übungen).

15.9 Definition: $(V, (\quad , \quad))$ sei euklidischer Vektorraum, $x, y \in V$ zwei Vektoren, die wir als Punkte des zu V gehörenden affinen Raumes auffassen. Dann ist der Abstand von x und y gegeben als

$$d(x,y) := \|x-y\| = (x-y, x-y)^{\frac{1}{2}}$$

15.10 Folgerung: Für die Abstandsfunktion $d : V \times V \to \mathbb{R}$ gilt:

i) $d(x,y) \geq 0 \quad \forall x, y \in V$ mit $d(x,y) = 0$ genau, wenn $x = y$ ist.

ii) $d(x,y) = d(y,x)$ für alle $x, y \in V$

iii) $d(x,y) + d(y,z) \geq d(x,z)$ für alle $x, y, z \in V$ (die sogenannte Dreiecksungleichung)

15.11 Bemerkung: Eine Menge M versehen mit einer Abstandsfunktion $d : M \times M \to \mathbb{R}$, die (i), (ii) und (iii) erfüllt, ist ein sogenannter *metrischer Raum.*

15.12 Satz: $(V, (\quad , \quad))$ sei ein euklidischer endlich dimensionaler Vektorraum, $\{e_1, \ldots, e_n\}$ eine Basis von V. Dann ist das Skalarprodukt $(\quad , \quad)$ eindeutig festgelegt durch die Werte

$$g_{ij} := (e_i, e_j) \quad (1 \leq i, j \leq n)$$

Beweis: Seien $x = \sum_{i=1}^{n} \xi_i e_i$, $y = \sum_{j=1}^{n} \eta_j e_j$ zwei beliebige Vektoren aus V. Es ergibt sich

$$(x,y) = \sum_{i,j=1}^{n} \xi_i \eta_j (e_i, e_j) \qquad = \sum_{j,j=1}^{n} g_{ij} \xi_i \eta_j$$

womit die Behauptung sofort folgt. □

15.13 Satz: V sei ein endlich dimensionaler Vektorraum. Dann ist die lineare Abbildung $\varphi : V \to V^\star$

$$\begin{cases} V & \to & V^\star \\ x & \mapsto & (\varphi_x : \quad V \quad \to \quad \mathbb{R}) \end{cases}, \text{ dabei } \varphi_x : y \mapsto (x,y)$$

ein Isomorphismus von V mit $V^\star$.

Beweis: Die Abbildung ist injektiv, denn beispielsweise ist $\varphi_x(x) = (x,x) = 0 \Leftrightarrow x = 0$. Wegen $\dim(V) = \dim(V^\star)$ ist die Abbildung dann auch surjektiv und daher ein Isomorphismus. □

15.14 Definition: Zwei Vektoren $x, y \in V$ eines euklidischen Vektorraums $(V, (\quad,\quad))$ sind orthogonal zu einander (stehen senkrecht aufeinander) genau dann, wenn gilt

$$(x,y) = 0$$

15.15 Definition: Sei $(V, (\quad,\quad))$ ein euklidischer Vektorraum.

i) Eine Basis $\{e_1, \ldots, e_n\}$ von V heißt *Orthogonalbasis* wenn gilt: $(e_i, e_j) = 0$, falls $i \neq j$.
Die Matrix $(g_{ij}) = ((e_i, e_j))$ ist dann eine Diagonalmatrix mit lauter positiven Zahlen auf der Diagonalen.

ii) $\{e_1, \ldots, e_n\}$ heißt *Orthonormalbasis*, wenn gilt:

$$(e_i, e_j) = \delta_{ij} \quad (1 \leq i, j \leq n)$$

Die Matrix $(g_{ij}) = ((e_i, e_j))$ ist dann die Einheitsmatrix.

15.16 Folgerung: Sei $(V, (\quad,\quad))$ ein euklidischer Vektorraum mit Orthogonalbasis $\{e_1, \ldots, e_n\}$. Dann gilt für jeden Vektor $x \in V$ die Entwicklung

$$x = \sum_{i=1}^{n} (x, e_i) e_i \quad .$$

Beweis: Sei also

$$x' := \sum_{i=1}^{n} (x, e_i) e_i \quad .$$

Es ist $(x, e_i) = (x', e_i)$ für $i = 1, \ldots, n$. Daraus folgt $(x - x', e_i) = 0$ für $i = 1, \ldots, n$. Dann ist aber $(x - x', y) = 0$ für alle $y \in V$, insbesondere $(x - x', x - x') = 0$. Daher ist $x - x' = 0$. $\square$

15.17 Das Orthonormalisierungsverfahren von E. Schmidt: Sei $(V, (\ \ ,\ \ ))$ ein euklidischer Vektorraum mit Basis $\{e_1, \ldots, e_n\}$. Wir konstruieren zu diesen Daten eine Orthonormalbasis $\{\tilde{e}_1, \ldots, \tilde{e}_n\}$ wie folgt:

$$\tilde{e}_1 := \frac{1}{\|e_1\|} e_1$$

Offenbar ist $\|\tilde{e}_1\| = 1$ bzw. $(\tilde{e}_1, \tilde{e}_1) = 1$ Seien $\{\tilde{e}_1, \ldots, \tilde{e}_i\}$ als orthonormales System bereits aus dem Teilsystem $\{e_1, \ldots, e_i\}$ konstruiert.

Insbesondere gelte

$$\mathbb{R}\tilde{e}_1 + \ldots + \mathbb{R}\tilde{e}_i = \mathbb{R}e_1 + \ldots + \mathbb{R}e_i$$

Wir setzen also voraus, dass auch $\{\tilde{e}_1, \ldots, \tilde{e}_i\}$ eine Basis des Vektorraums $\mathbb{R}e_1 + \ldots + \mathbb{R}e_i$ ist, die zusätzlich orthonormal ist.

Wir machen den Ansatz:

$$\tilde{\tilde{e}}_{i+1} = \lambda_1 \tilde{e}_1 + \ldots + \lambda_i \tilde{e}_i + e_{i+1}$$

Dann ergeben sich, falls $\{\tilde{e}_1, \ldots, \tilde{e}_{i+1}\}$ auch ein orthonormales System von Vektoren ist, die Bedingungen:

$$\left.\begin{array}{ccccc} 0 & = & (\tilde{e}_{i+1}, \tilde{e}_1) & = & \lambda_1 + (e_{i+1}, \tilde{e}_1) \\ \vdots & & & & \vdots \\ 0 & = & (\tilde{e}_{i+i}, \tilde{e}_i) & = & \lambda_i + (e_{i+1}, \tilde{e}_i) \end{array}\right\}$$

Daher setzen wir

$$\lambda_j := -(e_{i+1}, \tilde{e}_j) \quad (j = 1, \ldots, i)$$

und haben damit einen Vektor $\tilde{\tilde{e}}_{i+1}$ konstruiert, der senkrecht auf den Vektoren $\{\tilde{e}_1, \ldots, \tilde{e}_i\}$ steht. Weiter ist $\tilde{\tilde{e}}_{i+1} \neq 0$. Normieren wir diesen Vektor noch durch

$$\tilde{e}_{i+1} := \frac{1}{\|\tilde{\tilde{e}}_{i+1}\|} \tilde{\tilde{e}}_{i+1} \quad ,$$

so sieht man, dass das neue System $\{\tilde{e}_1, \ldots, \tilde{e}_{i+1}\}$ wieder ein Orthonormalsystem ist und außerdem wieder gilt:

$$\mathbb{R}\tilde{e}_1 + \ldots + \mathbb{R}\tilde{e}_{i+1} = \mathbb{R}e_1 + \ldots + \mathbb{R}e_{i+1}$$

Durch Iteration dieses Verfahrens ergibt sich die gesuchte Orthonormalbasis $\{\tilde{e}_1, \ldots, \tilde{e}_n\}$. Dies ist das sog. Orthonormalisierungsverfahren von E. Schmidt, einem Schüler von D. Hilbert.

15.18 Definition und Satz: $(V, (\quad,\quad))$ sei ein euklidischer Vektorraum, $W \subset V$ ein Untervektorraum.

i) W ist dann mit der Einschränkung des Skalarproduktes zu einer Abbildung $W \times W \to \mathbb{R}$ selber ein euklidischer Vektorraum.

ii) Bezeichnet
$$W^\perp := \{x \in V \mid (x,y) = 0 \text{ für alle } y \in W\}$$
das sog. orthogonale Komplement von W, so gilt: $W^\perp$ ist ein Untervektorraum von V und es ist
$$\dim(W^\perp) = \dim(V) - \dim(W)$$

iii) Es ist $((W)^\perp)^\perp = W$

Beweis:

i) ist klar.

ii) Seien $x_1, x_2 \in W^\perp$, dann ist wegen
$$(\lambda_1 x_1 + \lambda_2 x_2, y) = \lambda_1(x_1, y) + \lambda_2(x_2, y) = 0$$
für alle $y \in W$ auch $\lambda_1 x_1 + \lambda_2 x_2 \in W^\perp$ für alle $\lambda_1, \lambda_2 \in \mathbb{R}$.

Wir wählen ein Basis $\{e_1, \ldots, e_r\}$ von W (mit $r = \dim(W)$) und ergänzen sie zu einer Basis $\{e_1, \ldots, e_n\}$ von V.

Mit dem Orthogonalisierungsverfahren von E. Schmidt bestimmen wir zu $\{e_1, \ldots, e_n\}$ eine Orthonormalbasis $\{\tilde{e}_1, \ldots, \tilde{e}_n\}$, für die speziell gilt:
$$\mathbb{R}e_1 + \ldots + \mathbb{R}e_j = \mathbb{R}\tilde{e}_1 + \ldots + \mathbb{R}\tilde{e}_j \quad (j = 1, \ldots, n)$$
Insbesondere gilt:
$$W = \mathbb{R}e_1 + \ldots + \mathbb{R}e_r = \mathbb{R}\tilde{e}_1 + \ldots + \mathbb{R}\tilde{e}_r$$
Wegen der Orthonormalität der Basis $\{\tilde{e}_1, \ldots, \tilde{e}_n\}$ folgt dann aber sofort
$$W^\perp = \mathbb{R}\tilde{e}_{r+1} + \ldots + \mathbb{R}\tilde{e}_n.$$
Daher folgt
$$\dim(W^\perp) = n - \dim(W)$$

iii) Nach Definition folgt
$$(W^\perp)^\perp \supset W$$
Nach (ii) ist
$$\dim(W^\perp)^\perp = n - \dim(W^\perp) = n - (n - \dim(W)) = \dim(W)$$
Daher folgt für die obige Inklusion sogar Gleichheit: $(W^\perp)^\perp = W$. □

15.19 Folgerung:

i) Sei $(V, (\quad,\quad))$ ein euklidischer Vektorraum, $W \subset V$ ein beliebiger Teilraum. Dann hat man eine natürliche Zerlegung als direkte Summen

$$V = W \oplus W^{\perp}$$

ii) Speziell hat man für jedes $W \subset V$ eine Projektionsabbildung in $\mathrm{End}(V)$,

$$\mathrm{pr}_W : V \to W \quad ,$$

die jedem $x \in V$ seine Komponente in W zuordnet.

Beweis: Wir betrachten die lineare Abbildung

$$\left.\begin{array}{rccc} \varphi : & W \oplus W^{\perp} & \to & V \\ & (w, w') & \to & (w + w') \end{array}\right\}$$

φ ist injektiv denn anderenfalls wäre $\mathrm{Ker}(\varphi) \neq 0$ und es gäbe $(w, w') \in \mathrm{Ker}(\varphi)$ mit $w + w' = \varphi(w, w') = 0$.

Dann folgt $w = -w'$, d. h. $w \in W \cap W^{\perp}$. Daher ist $w = 0$ und ebenso $w' = 0$. Also ist φ injektiv und, da

$$\begin{aligned} \dim(W \oplus W^{\perp}) &= \dim(W) + \dim(W^{\perp}) \\ &= \dim(W) + (n - \dim(W)) \\ &= n = \dim(V) \end{aligned}$$

gilt, und surjektiv. Damit folgt (i).

Die Projektionsabbildung ergibt sich aus

$$V \xrightarrow{(\varphi^{-1})} W \oplus W^{\perp} \xrightarrow{\mathrm{pr}_1} W$$

durch Komposition und ist daher ebenfalls linear. $\square$

Am Ende dieses Abschnittes definieren wir noch einen weiteren grundlegenden Begriff der euklidischen Geometrie, den des Winkels.

15.20 Definition: Sei $(V, (\quad,\quad))$ ein euklidischer Vektorraum, $x, y \in V$ zwei Vektoren, $x, y \neq 0$. Dann wird der Winkel zwischen x und y, $\varphi = \sphericalangle(x, y)$, festgelegt durch

$$0 \leq \varphi \leq \pi \quad \text{und} \quad \cos(\varphi) = \frac{(x, y)}{\|x\| \cdot \|y\|} = \frac{(x, y)}{((x, x)(y, y))^{1/2}}$$

15.21 Bemerkung:

i) Nach den Gleichungen von Cauchy-Schwarz ist

$$-1 \leq \frac{(x,y)}{\|x\| \cdot \|y\|} \leq +1 \quad .$$

Da der Cosinus im Intervall $[0, \pi]$ eine stetige, streng monoton fallende Funktion ist, wird jeder Wert aus $[+1, -1]$ $(\cos(0) = +1, \cos(\pi) = -1$!) genau einmal angenommen. Daher ist der obige Winkel $\sphericalangle(x, y)$ eindeutig definiert.

ii) In der Elementargeometrie wurden Winkel anschaulich durch Kreisteilung des Vollwinkels von 360° eingeführt, systematischer kann man sie über die zugehörige Bogenlänge des Einheitskreises einführen. Will man es streng mathematisch machen, so braucht man in jeden Fall an dieser Stelle reelle Analysis.

15.22 Folgerung: Zwei Vektoren $x, y \in V$, sind orthogonal zueinander genau dann, wenn x oder y gleich Null sind oder $\sphericalangle(x, y) = \frac{\pi}{2}$ ist.

15.23 Die Hessesche Normalform: Gegeben sei im euklidischen Vektorraum $V \quad a \in V, \quad a \neq 0$. Die Gleichung

$$(a, x) = d \quad , \; d \in \mathbb{R} \quad ,$$

beschreibt eine Hyperebene, die in der Regel nicht durch den Nullpunkt geht. Die Gleichung

$$\left(\frac{a}{\|a\|}, x \right) = \frac{d}{\|a\|}$$

beschreibt dieselbe Hyperebene, weswegen wir ohne Beschränkung der Allgemeinheit annehmen können, dass $\|a\| = 1$ ist.

15.24 Satz: Für $x \in V$ beliebig gibt $|(a, x) - d|$ genau den Abstand von x von der Hyperebene $H : (a, x) = d$ an.

Beweis: Gesucht ist ein Vektor y, so dass $x + y \in H$ ist und $y \in (a)^{\perp}$, d. h. y steht senkrecht auf den Vektoren in H. Schreibe $x = (x, a)a + x'$.
Es folgt: $x' \in (a)^{\perp}$.
$((x, a) - d) \;\; a =: y$ liefert das Verlangte. □

15.25 Bemerkung: Anhand des Vorzeichens von $(x, a) - d$ kann man noch leicht feststellen, ob der Nullpunkt und der Punkt x auf der selben oder verschiedenen Seiten der Hyperebene H liegen.

16 Isometrien und orthogonale Gruppen

$(V, (\ \ ,\ \ ))$ und $(W, (\ \ ,\ \ ))$ seien euklidische Vektorräume. Wie immer sind wir an den Abbildungen besonders interessiert, die die Struktur erhalten.

16.1 Definition: Eine lineare Abbildung $\varphi : V \to W$ heißt Isometrie $:\Leftrightarrow \forall x, y \in V$ gilt:

$$(\varphi(x), \varphi(y)) = (x, y)$$

16.2 Satz:

i) Eine Isometrie φ ist stets injektiv.

ii) φ erhält Längen und Winkel.

iii) Ist φ bijektiv, so ist auch φ^{-1} eine Isometrie.

Beweis:

ii) $||x|| = (x, x)^{\frac{1}{2}} = (\varphi(x), \varphi(x))^{\frac{1}{2}} = ||\varphi(x)||$ für alle $x \in V$.
Entsprechend: Ist $\alpha = \angle(x, y)$, so gilt

$$\cos(\alpha) = \frac{(x, y)}{||x||||y||}$$

Dann ergibt sich sofort für $\beta = \angle(\varphi(x), \varphi(y))$:

$$\cos(\beta) = \frac{(\varphi(x), \varphi(y))}{||\varphi(x)||||\varphi(y)||} = \frac{(x, y)}{||x||||y||} = \cos(\alpha)$$

Wegen $0 \leq \alpha, \beta < \pi$ folgt: $\alpha = \beta$, wegen der strengen Monotonie des Cosinus in diesem Bereich.

i) Wäre $\varphi(x) = 0$, so folgte $0 = ||\varphi(x)|| = ||x||$. Daher $x = 0$ und damit die Injektivität.

iii) Sicherlich ist φ^{-1} linear. Sind $u, v \in W$, so schreibe $u = \varphi(x), v = \varphi(y)$ und damit

$$(u, v) = (\varphi(x), \varphi(y)) = (x, y) = (\varphi^{-1}(u), \varphi^{-1}(v))$$

Also ist φ^{-1} ebenfalls eine Isometrie.

16.3 Proposition: Sind $\varphi : U \to V, \psi : V \to W$ Isometrien der euklidischen Vektorräume $(U, (\ \ ,\ \ )), (V, (\ \ ,\ \ ))$ und $(W, (\ \ ,\ \ ))$, so ist auch die zusammengesetzte Abbildung $\psi \circ \varphi : U \to W$ eine Isometrie.

Beweis: Übung □

Bemerkung: Ist $\varphi : V \to W$ eine Isometrie und ist $\dim(V) = \dim(W) < \infty$, so ist φ bijektiv. Dies ist klar, denn aus der Injektivität von φ folgt unter diesen Umständen die Surjektivität.

16.4 Definition und Satz: Sei $(V, (\ \ ,\ \ ))$ ein euklidischer Vektorraum. Die Menge aller bijektiven Isometrien $\varphi : V \to V$ bildet dann die sog. orthogonale Gruppe von $(V, (\ \ ,\ \ ))$: Schreibweise $O(V, (\ \ ,\ \ ))$ oder einfacher $O(V)$.

Bemerkung: Statt von Isometrien spricht man auch von orthogonalen Abbildungen.

16.5 Beispiele orthogonaler Abbildungen:

i) $V := \mathbb{R}^2$ mit Standardmetrik: $x = (\xi_1, \xi_2), y = (\eta_1, \eta_2)$ und $x \cdot y = \xi_1\eta_1 + \xi_2\eta_2$.
Die Menge der Drehungen

$$\left\{ \begin{pmatrix} \cos\alpha, & -\sin\alpha \\ \sin\alpha, & \cos\alpha \end{pmatrix} \middle| \, 0 \leq \alpha < 2\pi \right\}$$

bildet eine Untergruppe der orthogonalen Gruppe $O(V)$. Tatsächlich handelt es sich um die sog. spezielle orthogonale Gruppe.

ii) Sei $(V, (\ \ ,\ \ ))$ beliebiger euklidischer Vektorraum, $a \in V$, $a \neq 0$.
Dann ist die Abbildung

$$S_a : V \to V, x \mapsto x - \frac{2(a,x)}{(a,a)} a$$

eine sog. Spiegelung (an der zu a orthogonalen Hyperebene)

$$H_a := \{x \in V \mid (a, x) = 0\}$$

Beweis: Die Linearität und Isometrieeigenschaft rechnet man unmittelbar nach. Es ist weiter $S_a(x) = x$ für alle $x \in H_a, S_a(a) = -a$. Dies erklärt den Begriff „Spiegelung an H_a“.

16.6 Satz: Sei $\{e_1, \dots, e_n\}$ eine Basis des euklidischen Vektorraums V. $\varphi : V \to W$ ist eine Isometrie in den euklidischen Vektorraum W $\Leftrightarrow$ Es gilt $(\varphi(e_i), \varphi(e_j)) = (e_i, e_j)$ für $1 \leq i, j \leq n$.

Beweis: Sei $x = \sum_{i=1}^n \xi_i e_i, \sum_{j=1}^n \eta_j e_j$. Dann ist also

$$\begin{aligned}
(\varphi(x), \varphi(y)) &= (\sum_{i=1}^n \xi_i \varphi(e_i), \sum_{j=1}^n \eta_j \varphi(e_j)) \\
&= \sum_{i,j=1}^n \xi_i \eta_j (\varphi(e_i), \varphi(e_j)) \\
&= \sum_{i,j=1}^n \xi_i \eta_j (e_i, e_j) \\
&= (\sum_{i=1}^n \xi_i e_i, \sum_{j=1}^n \eta_j e_j) \\
&= (x, y)
\end{aligned}$$

Dies zeigt eine Richtung, die andere ist trivial. $\square$

16.7 Folgerung: V, W seien euklidische Vektorräume, $\{e_1, \ldots, e_n\}$ eine beliebige Orthonormalbasis von V, $\{f_1, \ldots, f_n\}$ eine beliebige Orthonormalbasis von W. Dann gibt es genau eine Isometrie $\varphi : V \to W, \varphi(e_i) = f_i$ für $i = 1, \ldots, n$.

Sei jetzt wieder V n-dimensionaler euklidischer Vektorraum, $\{e_1, \ldots, e_n\}$ eine beliebige Basis von V, $G = ((e_i, e_j))_{1 \leq i,j \leq n}$ der „metrische Tensor“ von V. $\varphi : V \to V$ sei eine beliebige Isometrie von V auf sich, es sei $A = Matrix(\varphi : \{e_1, \ldots, e_n\})$ die zugehörige Matrix.
Also gilt

$$\varphi(e_j) = \sum_{i=1}^n \alpha_{ij} e_i$$

Die Bedingung der Orthogonalität von φ bedeutet also: es ist

$$\begin{aligned}
g_{kl} &= (\varphi(e_k), \varphi(e_l)) \\
&= (\sum_{i=1}^n \alpha_{ik} e_i, \sum_{j=1}^n \alpha_{jl} e_j) \\
&= \sum_{i=1}^n \sum_{j=1}^n \alpha_{ik} (e_i, e_j) \alpha_{jl} \\
&= \sum_{i=1}^n \sum_{j=1}^n \alpha_{ik} g_{ij} \alpha_{jl}
\end{aligned}$$

Man rechnet leicht nach, dass dies in Matrixschreibweise die Identität $A^t G A = G$ bedeutet.

16.8 Definition und Satz: Eine Matrix A ist orthogonal bzgl. der die euklidische Metrik festlegenden symmetrischen Matrix G genau dann, wenn gilt:

$$A^t G A = G$$

Bemerkungen:

i) Zum Beispiel rechnet man jetzt die Gruppeneigenschaft der orthogonalen Matrizen leicht nach:

Gilt
$$A^tGA = G \quad \text{sowie} \quad B^tGB = G$$
so folgt für das Produkt (AB) sofort:
$$(AB)^tG(AB) = B^t(A^tGA)B = B^tGB = G$$

ii) Besonders wichtig ist der Fall $G = E$, wo also die Basis $\{e_1, \ldots, e_n\}$ von V eine Orthonormalbasis ist. Dann ist also $G = E$ und es gilt:

16.9 Folgerung: A ist eine orthogonale Matrix $\Leftrightarrow$ Es gilt $A^tA = E$
oder äquivalent: $A^t = A^{-1}$.

iii) Orthogonale Matrizen A dieser Art haben, wie man sich sofort überlegt, folgende Eigenschaften: Bezüglich des gewöhnlichen euklidischen Skalarproduktes haben alle Spalten (Zeilen) von A die Länge 1; das Skalarprodukt verschiedener Spalten (Zeilen) ist Null. Für die Spalten folgt die Aussage nämlich, weil die Spalten die Bilder der Basisvektoren sind und die Basis eine Orthonormalbasis war. Da mit A auch $A^t = A^{-1}$ eine orthogonale Matrix bzgl. $G = E$ ist, folgt durch Betrachtung von A^t die entsprechende Aussage für die Zeilen.

16.10 Satz:

i) Die reellen Eigenwerte einer Isometrie $\varphi : V \to W$ sind ± 1 (sofern überhaupt reelle Eigenwerte vorhanden sind).

ii) Die Determinante einer orthogonalen Abbildung ist ± 1.

Beweis:

i) Sei $\varphi(x) = \lambda x$ mit $x \in V$, $x \neq 0$ und $\lambda \in \mathbb{R}$. Dann folgt
$$\begin{aligned} (\varphi(x), \varphi(x)) &= (x, x) \quad \text{, aber auch} \\ (\varphi(x), \varphi(x)) &= (\lambda x, \lambda x) = \lambda^2 (x, x) \end{aligned}$$
Es folgt wegen $(x, x) \neq 0$ daher $\lambda^2 = 1$, also $\lambda = \pm 1$.

ii) Bezüglich einer Basis $\{e_1, \ldots, e_n\}$ gilt für eine orthogonale Matrix A
$$A^tGA = G$$
Wegen $\det(G) \neq 0$ und $\det(A^t) = \det(A)$ folgt $\det(A)^2 = 1$ und damit (ii). $\square$

Bemerkung:

i) Wie die Rotationen $\begin{pmatrix} \cos\alpha, & -\sin\alpha \\ \sin\alpha, & \cos\alpha \end{pmatrix}$ im $\mathbb{R}^2$ zeigen, gibt es orthogonale Matrizen ohne reelle Eigenwerte.

ii) Eine orthogonale Abbildung im $\mathbb{R}^3$ hat immer wenigstens einen reellen Eigenwert. Sei dieser $+1$ und e_1 ein zugehöriger Eigenvektor, also $\varphi(e_1) = e_1$, so folgt

$$\varphi(e_1)^\perp = (e_1)^\perp$$

Auf $(e_1)^\perp$ ist φ dann bei Einschränkung entweder eine Drehung oder eine Spiegelung (für eine nähere Diskussion siehe die Übungen).

16.11 Definition und Satz: Sei V ein endlichdimensionaler euklidischer Vektorraum. Es ist

$$SO(V) = \{\varphi : V \to V \mid \varphi \in O(V) \text{ und } \det(\varphi) = 1\}$$

die sog. spezielle orthogonale Gruppe. Diese ist ein Normalteiler in der reellen orthogonalen Gruppe $O(V)$. Genauer ist sie der Kern des Determinantenhomomorphismus

$$\det : O(V) \to (\mathbb{R}^\times, \cdot) \quad \text{bzw.} \quad \det : O(V) \to (\pm 1, \cdot)$$

Beweis: Übung. $\square$

Zuletzt noch zwei schöne geometrische Sätze:

16.12 Satz: Jede orthogonale Abbildung $\varphi : V \to V$ eines n-dimensionalen euklidischen Raumes ist Produkt von höchstens n Spiegelungen.

Beweis: Sei $e \in V, e \neq 0$ ein beliebiger Vektor. Wir finden eine Spiegelung S mit $S_1(e) = \varphi(e)$. Daraus folgt: Es ist

$$S_1^{-1}\varphi(e) = e$$

Dann folgt $(S_1^{-1}\varphi)(e)^\perp = (e)^\perp$. Da $\dim(e)^\perp = \dim(V) - 1$, können wir $S_1^{-1}\varphi|_{(e)^\perp} = \tilde{S}_2 \ldots \tilde{S}_r$ schreiben mit $r \leq n$. Jedes $\tilde{S}_i$ ist Spiegelung auf $(e)^\perp$ und wird durch die

Definition:

$$\begin{cases} S_i(x) & := \tilde{S}_i(x) \quad \text{für} \quad x \in (e)^\perp \\ S_i(e) & := e \end{cases}$$

Zu einer Spiegelung auf ganz V fortgesetzt. Es gilt

$$S_1^{-1}\varphi = S_2 \ldots S_r,$$

denn es gilt

$$S_1^{-1}\varphi|_{(e_1)^\perp} = (S_2 \ldots S_r)|_{(e_1)^\perp}$$
$$S_1^{-1}\varphi(e_1) = (S_2 \ldots S_r)e_1 = e_1$$

Dann folgt aber:

$$\varphi = S_1 S_2 \dots S_r$$

□

16.13 Satz: Sei $\varphi : V \to V$ eine orthogonale Abbildung. Dann gibt es eine Orthonormalbasis von V, bzgl. der φ sich in Blöcken der Form

$$\mathbb{M}(\varphi) = \begin{pmatrix} A_1 & & 0 \\ & \ddots & \\ 0 & & A_r \end{pmatrix}$$

schreiben lässt, dabei sind die A_j

i) entweder 1-Blöcke, $A_j = \pm 1$ oder -1, also eine 1×1-Matrix oder es ist

ii) A_j eine 2×2-Matrix der Form

$$\begin{pmatrix} \cos \alpha_j, & -\sin \alpha_j \\ \sin \alpha_j, & \cos \alpha_j \end{pmatrix}, 0 \leq \alpha_j < \pi$$

also eine Drehmatrix. Diese Darstellung von φ ist im Wesentlichen eindeutig.

Beweis: Ohne Einschränkung denken wir uns φ durch eine reelle Matrix A als Abbildung

$$A : \mathbb{R}^n \to \mathbb{R}^n, x \mapsto A \cdot x$$

gegeben. Genausogut können wir dieses A aber auch als Abbildung

$$A : \mathbb{C}^n \to \mathbb{C}^n, x \mapsto A \cdot x$$

auffassen. Sei $e_1 \in \mathbb{C}^n$ ein beliebiger Eigenvektor zum Eigenwert $\lambda_1 \in \mathbb{C}$. Ist dabei sogar $\lambda_1 \in \mathbb{R}$, so können wir sogar annehmen, dass $e_1 \in V \simeq \mathbb{R}^n$ reell ist. Es folgt: $\lambda_1 = \pm 1$. Außerdem respektiert φ die orthogonale Zerlegung. Auf $(\mathbb{R}e_1)^\perp$ können wir durch geeignete Basiswahl φ in der gewünschten Form annehmen. Damit folgt die Behauptung in diesem Fall.
Sei jetzt $\lambda_1 \in \mathbb{C}, \lambda_1 \notin \mathbb{R}$. Dann ist auch ein zugehöriger Eigenvektor
$e_1 = (\xi_1, \dots, \xi_n)$ in $\mathbb{C}^n, \notin \mathbb{R}^n = V$. Es gilt also

$$Ae_1 = \lambda_1 e_1$$

Mit $\bar{e_1} := (\bar{\xi_1}, \dots, \bar{\xi_n}) \in \mathbb{C}^n$ gilt dann aber auch

$$A\bar{e_1} = \bar{\lambda_1}\bar{e_1}$$

Dann rechnet man sofort nach: $f_1 := e_1 + \bar{e_1} \in V = \mathbb{R}^n$, genauso ist $f_2 := \frac{1}{2i}(e_1 - \bar{e_1}) \in V = \mathbb{R}^n$.

Weiter gilt für

$$W := \mathbb{R}f_1 + \mathbb{R}f_2 \subset V = \mathbb{R}$$

$A(W) = W$ und es ist A orthogonal sowie $\det(A) = \lambda, \bar{\lambda}$ (Beweis hierfür!) positiv reell. Also ist $A|_{W_1}$ eine Drehung um einen Winkel α_1 mit $0 \leq \alpha_1 < \pi$. Da A weiter nach Induktionsannahme auf $W_1^\perp$ bereits in obiger Form angenommen werden kann, folgt die Behauptung. □

17 Hauptachsentransformation, selbstadjungierte Abbildungen

Gegeben sei der euklidische $\mathbb{R}^2$ oder $\mathbb{R}^3$ mit Standardmetrik, in ihm eine Ellipse oder ein Ellipsoid E mit Mittelpunkt $0 = (0,0) \in \mathbb{R}^2$ beziehungsweise $0 = (0,0,0) \in \mathbb{R}^3$, so dass mit $x \in \mathbb{R}^2$ beziehungsweise $\mathbb{R}^3$, wenn es auf E liegt, auch $(-x)$ auf E liegt. Wie man sich geometrisch anschaulich leicht überlegt, gibt es dann stets eine kürzeste und eine längste Achse der Ellipse oder des Ellipsoides, die zusätzlich auf einander senkrecht stehen. Diesen wichtigen geometrischen Sachverhalt gilt es besser zu verstehen.

Gegeben sei also ein endlichdimensionaler Vektorraum V, $\varphi : V \to V$ sei ein Endomorphismus (Operator).

17.1 Definition: Die Abbildung φ heißt selbstadjungiert $:\Longleftrightarrow$ Für alle $x, y \in V$ gilt:

$$(x, \varphi(y)) = (\varphi(x), y)$$

Für selbstadjungierte lineare Abbildungen gilt

17.2 Satz: Ist $e \in V$ ein Eigenvektor von φ zum Eigenwert λ, so ist $(\mathbb{R}e)^\perp$ ebenfalls invariant unter φ, d. h. es gilt $\varphi(\mathbb{R}e)^\perp \subset (\mathbb{R}e)^\perp$.

Beweis: Sei $x \in (\mathbb{R}e)^\perp$. Es gilt $(\varphi(x), e) = (x, \varphi(e))$, da φ selbstadjungiert ist. Dann ist aber

$$(x, \varphi(e)) = (x, \lambda e) = \lambda(x, e) = 0$$

wegen $x \in (\mathbb{R}e)^\perp$. Also folgt auch $\varphi(x) \in (\mathbb{R}e)^\perp$. $\square$

17.3 Satz: Jede selbstadjungierte Abbildung hat wenigstens einen reellen Eigenwert.

Beweis: O.E.d.A. können wir annehmen: $V = \mathbb{R}^n$ mit dem Standardskalarprodukt versehen. Die Abbildung φ ist gegeben durch eine Matrix

$$A : \mathbb{R}^n \to \mathbb{R}^n \ , \quad x \mapsto A \cdot x$$

Wir erweitern zu $V_\mathbb{C} = \mathbb{C}^n$ mit hermiteschem Skalarprodukt gegeben durch

$$(x, y) = \sum_{i=1}^{n} \xi_i \overline{\xi_i} \quad \text{dabei}$$

$$x = (\xi_1, \ldots, \xi_n) \in \mathbb{C}^n \ , \quad y = (\xi_1, \ldots, \xi_n) \in \mathbb{C}^n \ .$$

A definiert weiterhin eine lineare Abbildung

$$A : \mathbb{C}^n \to \mathbb{C}^n \ , \quad x \mapsto A \cdot x$$

Wie man sofort nachrechnet, gilt nach wie vor Selbstadjungiertheit, d. h.

$$(Ax, y) = (x, Ay) \ .$$

Sei $\tilde{e} \in \mathbb{C}^n$ Eigenvektor zum Eigenwert $\lambda \in \mathbb{C}$. Es folgt:

$$\begin{aligned} (A\tilde{e}, \tilde{e}) &= (\lambda\tilde{e}, \tilde{e}) = \lambda(\tilde{e}, \tilde{e}) \ , \quad \text{andererseits} \\ (A\tilde{e}, \tilde{e}) &= (\tilde{e}, A\tilde{e}) = (\tilde{e}, \lambda\tilde{e}) = \overline{\lambda}(\tilde{e}, \tilde{e}) \end{aligned}$$

Wegen $(\tilde{e}, \tilde{e}) \neq 0$ folgt: $\lambda = \overline{\lambda}$, d. h. λ ist automatisch reell.

Dann gibt es aber zum reellen Eigenwert λ einen reellen Eigenvektor $e \in \mathbb{R}^n$, der eventuell verschieden ist von $\tilde{e}$. Die Behauptung folgt. □

17.4 Hauptsatz über selbstadjungierte Abbildungen: Sei $(V, (\ \ ,\ \ ))$ ein endlichdimensionaler euklidischer Vektorraum, $\varphi : V \to V$ eine bezüglich $(\ \ ,\ \ )$ selbstadjungierte Abbildung. Dann gibt es eine Orthonormalbasis $\{e_1, e_2, \ldots, e_n\}$ von V aus lauter Eigenvektoren von φ. φ hat die Gestalt

$$\mathbb{M}\left(\varphi : \{e_1, \ldots, e_n\}\right) = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$$

mit lauter reellen Eigenwerten $\lambda_1, \ldots, \lambda_n$.

Beweis: (Durch vollständige Induktion nach $\dim(V) = n$)

i) $n = 1$: Es ist nichts zu beweisen.

ii) *Induktionsschritt:* Nach Satz **17.3** finden wir $e_1 \in V$, $\varphi(e_1) = \lambda_1 e_1$ mit $\lambda_1 \in \mathbb{R}$. O.E.d.A können wir dabei durch Normierung annehmen, dass $(e_1, e_1) = 1$ ist.

   Da φ selbstadjungiert ist, folgt

   $$\varphi\left((\mathbb{R}e_1)^\perp\right) \subseteq (\mathbb{R}e_1)^\perp$$

   Wegen $\dim(\mathbb{R}e_1)^\perp = \dim(V) - 1$ können wir die Induktionsannahme verwenden und eine Orthonormalbasis $\{e_2, \ldots, e_n\}$ von $(\mathbb{R}e_1)^\perp$ finden, in der $\varphi|(\mathbb{R}e_1)^\perp$ in der gewünschten Form vorliegt. $\{e_1, \ldots, e_n\}$ kann als gesuchte Basis von V genommen werden. □

Wir wenden das Ergebnis jetzt auf das eingangs dieses Kapitels beschriebene geometrische Problem an.

17.5 Satz: (Hauptachsentransformation) $(V, (\quad,\quad))$ sei ein endlich dimensionaler euklidischer Vektorraum, B sei ein weiteres Skalarprodukt auf V.

$$B : V \times V \to \mathbb{R},\ (x,y) \mapsto B(x,y).$$

Dann gibt es eine Orthonormalbasis von V, die auch für B noch eine Orthogonalbasis ist.

Beweis: Wir finden eine lineare Abbildung $\varphi : V \to V$, so dass gilt:

$$B(x,y) = (\varphi(x), y) \qquad \text{für alle } x, y \in V\ .$$

Dies werden wir unten begründen. φ ist selbstadjungt, denn es gilt $(x, \varphi(y)) = (\varphi(y), x) = B(y,x) = B(x,y)$, daher folgt:

$$(\varphi(x), y) = (x, \varphi(y)) \qquad .$$

Mit dem Hauptsatz über selbstadjungierte lineare Abbildungen folgt die Existenz einer Orthonormalbasis $\{e_1, \ldots, e_n\}$ von V aus Eigenwerten für φ.

Dann gilt aber

$$\begin{aligned} B(e_i, e_j) &= (\varphi e_i, e_j) \\ &= (\lambda_i e_i, e_j) \\ &= 0 \qquad \text{falls } i \neq j \text{ ist.} \end{aligned}$$

$\{e_1, \ldots, e_n\}$ ist also eine Orthogonalbasis für die Form B.

Es bleibt die Existenz von φ zu zeigen:
Zum Beispiel durch explizite Rechnung sieht man das so:
O.E.d.A. $V = \mathbb{R}^n, (\quad,\quad) =$ Standardskalarprodukt wie oben.
$B(x,y) = (xBy^t)$ mit $x = (\xi_1, \ldots, \xi_n)$, $y = (\eta_1, \ldots, \eta_n)$. Dann sei φ gegeben durch die Matrix B selbst, $B : \mathbb{R}^n \to \mathbb{R}^n$. Es ergibt sich nämlich:

$$\begin{aligned} (\varphi x, y) &= \left(Bx^t\right)^t \cdot y^t \\ &= xB^t y^t \\ &= xBy^t \end{aligned}$$

□

17.6 Bemerkungen:

a) Tatsächlich haben wir im Beweis nicht benutzt, dass B positiv definit ist. Die Bilinearität und Symmetrie von B hätte gereicht!

b) Bezüglich der gefundenden Basis $\{e_1, \ldots, e_n\}$ ist also mit $x = (\xi_1, \ldots, \xi_n)$

$$\begin{aligned} (x,x) &= \sum_{i=1}^n \xi_i^2 \\ B(x,x) &= \sum_{i=1}^n \lambda_i \xi_i^2 \end{aligned}$$

$\{e_1, \ldots, e_n\}$ bilden gerade die Hauptachsen.

Die Einheitssphäre bezüglich B wird also gegeben durch $\sum_{i=1}^{n} \lambda_i \xi_i^2 = 1$, dabei $\lambda_i > 0$: Wir nehmen $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n > 0$ an o.E.d.A. Für die Größe der Hauptachsen des so gegebenen Ellipsoids erhalten wir also

$$\left(\sqrt{\lambda_1}\right)^{-1} \leq \cdots \leq \left(\sqrt{\lambda_n}\right)^{-1}$$

c) Natürlich hätten wir wieder den Inhalt dieses gesamten Abschnittes für den Fall unitärer (oder auch hermitescher) Räume duchführen können. Es wäre alles wörtlich genauso gut gegangen.

Normale Endomorphismen: Man kann die Theorie dieses Paragraphen noch etwas erweitern. Es sei hierzu $(V, (\quad,\quad))$ ein *unitärer Vektorraum*.

17.7 Definition: Ein Endomorphismus $\varphi : V \to V$ heißt normal $:\Leftrightarrow \varphi$ und die adjungierte Abbildung φ^t (bezüglich des hermiteschen Produkts $(,)$) kommutieren miteinander, d. h. es gilt: $\varphi\varphi^t = \varphi^t\varphi$.

Es gilt wiederum:

17.8 Satz: Es gibt eine Orthonormalbasis $\{e_1, \ldots, e_n\}$ von $(V, (\quad,\quad))$ aus Eigenvektoren von φ. Umgekehrt ist jeder derartigen Endomorphismus φ, für den es eine solche Orthonormalbasis gibt, ein normaler Endomorphismus.

Beweis: Der zweite Teil des Satzes ist klar. Den ersten Teil sieht man so: Sicherlich gibt es einen Eigenvektor $e_1 \in V$ mit Eigenwert λ, d.h. $\varphi(e_1) = \lambda_1 e_1$. Dann folgt aber auch: $\varphi^t(e_1) = \overline{\lambda_1} e_1$, denn es ist allgemein für einen normalen Endomorphismus Ψ

$$\begin{aligned}
(\Psi x, \Psi x) &= \left(x, \Psi^t \Psi x\right) \\
&= \left(x, \Psi \Psi^t x\right) \quad (\text{wegen } \Psi\Psi^t = \Psi^t\Psi) \\
&= \left(\Psi^t x, \Psi^t x\right)
\end{aligned}$$

Dies angewendet auf $\Psi := \varphi - \lambda id$ und $\Psi^t = \varphi^t - \overline{\lambda} id$ folgt für einen Eigenvektor e_1, $\varphi e_1 = \lambda_1 e_1$ sofort:

$$0 = ((\varphi - \lambda_1)\, e_1, (\varphi - \lambda_1)\, e_1) = \left(\left(\varphi^t - \overline{\lambda_1}\right) e_i, \left(\varphi^t - \overline{\lambda_1}\right) e_1\right)$$

Daher aber sofort $\left(\varphi^t - \overline{\lambda_1}\right)(e_1) = 0$, d. h. $\varphi^t(e_1) = \overline{\lambda_1} e_1$

Hieraus folgt weiter sofort: $\varphi\left((\mathbb{C} \cdot e_1)^\perp\right) \subseteq (\mathbb{C}e_1)^\perp$, denn sei $x \in (\mathbb{C}e_1)^\perp$, so folgt sofort:

$$\begin{aligned}
(e_1, \varphi(x)) &= \left(\varphi^t(e_1), x\right) \\
&= \left(\overline{\lambda}, e_1, x\right) \\
&= \overline{\lambda_1}(e_1, x) \\
&= 0
\end{aligned}$$

Daher also $\varphi(x) \in (\mathbb{C}e_1)^\perp$. Danach verläuft der Beweis völlig entsprechend dem Fall selbstadjungierter Endomorphismen. $\square$

18 Ergänzungen

A Volumina. Eine Formel für die Gramsche Matrix.

Die gesamte Determinatentheorie war motiviert durch den Begriff des orientierten Volumens. Es ist daher nicht verwunderlich, wenn wir folgende Definition treffen.

18.1 Definition: Sei V ein n-dimensionaler reeller Vektorraum, $\{e_1, \ldots, e_n\}$ eine fixierte Basis, $\{a_1, \ldots, a_n\} \subset V$ ein geordnetes System von n Vektoren. Dem Parallelkörper

$$X(a_1, \ldots, a_n) := \left\{ \sum_{i=1}^{n} \xi_i a_i | 0 \le \xi_i \le 1 \right\}$$

ist dann das orientierte Volumen $\det(\alpha_{ij})_{1 \le i,j \le n}$ zugeordnet, dabei ist die Matrix $(\alpha_{ij})_{1 \le i,j \le n}$ gegeben durch

$$a_j = \sum_{i=1}^{n} \alpha_{ij} e_i \quad (j = 1, \ldots, n) \quad .$$

Das Volumen wird definiert als $\mathrm{vol}(X) := |\det((\alpha_{ij})_{1 \le i,j \le n})|$.

Gramsche Matrix: Sei $\{a_1, \ldots, a_n\}$ wie oben, $A = (\alpha_{ij})$, dabei wieder

$$a_j = \sum_{i=1}^{n} \alpha_{ij} e_i \quad .$$

Sei jetzt vorausgesetzt, dass V euklidisch ist und $\{e_i\}$ eine Orthonormalbasis. Dann ergibt sich für die Gramsche Matrix

$$G := ((a_i, a_j))_{1 \le i,j \le n} = (e_i (A^t A) e_j) \quad .$$

Speziell folgt noch der wichtige Zusammenhang

18.2 Satz:

i) Für die Gramsche Matrix gilt $G = A^t A$.

ii) $\det(G) = \det(A)^2$.

B Orientierung

Seien $\{e_1, \ldots, e_n\}$ und $\{f_1, \ldots, f_n\}$ zwei beliebige Basen eines reellen Vektorraumes.

18.3 Definition: Die Basen $\{e_1, \ldots, e_n\}$ und $\{f_1, \ldots, f_n\}$ geben dieselbe Orientierung des Vektorraumes V :$\Leftrightarrow$ Mit $\varphi(e_i) = f_i \quad (i = 1, \ldots, n)$ wird eine lineare Abbildung $\varphi : V \to V$ festgelegt und es ist $\det(\varphi) > 0$.

18.4 Satz: Die Relation „gleiche Orientierung" ist eine Äquivalenzrelation auf der Menge der Basen des reellen Vektorraumes V.

„Gleiche Orientierung" ist ein topologischer Begriff, invariant gegenüber stetiger Deformation. Genauer:

18.5 Satz: Sei $V = \mathbb{R}^n$, $[a, b] \subset \mathbb{R}$ ein Intervall, $\{a_1(t), \ldots, a_n(t)\}$ für alle $t \in [a, b]$ eine Basis, so dass mit $a_i(t) := (\alpha_{i1}(t), \ldots, \alpha_{in}(t)) \quad (i = 1, \ldots, n)$ alle Funktionen $\alpha_{ij}(t)$, $\alpha_{ij} : I \to \mathbb{R}$ stetige Funktionen sind. Dann definieren die Basen $\{a_1(t), \ldots, a_n(t)\}$ dieselbe Orientierung auf $V = \mathbb{R}^n$ für alle $t \in I$.

Beweis: Mit Hilfe der Entwicklungsformel folgt sofort:

$$\begin{array}{rccc} f: & [a,b] & \to & \mathbb{R} \\ & t & \mapsto & \det(\alpha_{ij}(t)) \end{array}$$

ist eine stetige Funktion auf $[a, b]$. Wegen des Zwischenwertsatzes folgt dann:

$$\begin{array}{lll} \text{Entweder} & \det(\alpha_{ij}(t)) > 0 & \text{für alle } t \in [a,b] \\ \text{oder} & \det(\alpha_{ij}(t)) < 0 & \text{für alle } t \in [a,b] \end{array}$$

□

C Das Vektorprodukt

Gegeben sei ein dreidimensionaler, euklidischer Vektorraum V mit Skalarprodukt $(\quad,\quad)$ und ausgezeichneter Orthonormalbasis $\{e_1, e_2, e_3\}$, etwa der $\mathbb{R}^3$ mit Standardskalarprodukt sowie der Standardbasis $\{e_1, e_2, e_3\}$.

Anschauliche Beschreibung des Vektorproduktes: Gegeben $a, b \in V$, dann ist

- $a \times b$ senkrecht auf a, b,
- und falls a und b linear unabhängig, ist $\{a, b, a \times b\}$ positiv orientiert (d.h. gleich orientiert wie $\{e_1, e_2, e_3\}$,
- es ist $|a \times b| =$ Flächeninhalt $X(a, b)$, dem durch a und b gegebene Parallelogramm.

Man stellt mit dieser Beschreibung leicht fest:

i) $V \times V \to V$, $(a, b) \mapsto a \times b$ ist bilinear in a, b.

ii) Es ist $e_1 \times e_2 = e_3$, $e_3 \times e_1 = e_2$ und $e_2 \times e_3 = e_1$ (zyklische Vertauschung der Indizies), entsprechend $e_2 \times e_1 = -e_3$, $e_3 \times e_2 = -e_1$, $e_1 \times e_3 = -e_2$.

Mit $a = \alpha_1 e_1 + \alpha_2 e_2 + \alpha_3 e_3$ und $b = \beta_1 e_1 + \beta_2 e_2 + \beta_3 e_3$ ergeben sich sofort die Formeln:

$$(a \times b) = \left.\begin{array}{rl} & (\alpha_2\beta_3 - \alpha_3\beta_2)e_1 \\ + & (\alpha_3\beta_1 - \alpha_1\beta_3)e_2 \\ + & (\alpha_1\beta_2 - \alpha_2\beta_1)e_3 \end{array}\right\}$$

Es gilt: $(a \times b, c) = \det(a, b, c)$.

Beweis: Beide Seiten sind linear in a, b und c. Weiter gilt die Gleichheit für die Fälle $a, b, c \in \{e_1, e_2, e_3\}$. Damit folgt leicht die Behauptung. $\square$

18.6 Proposition: Es gilt

i) $(a \times b) \perp a, b$

ii) $||a \times b|| = \mathrm{vol}\langle a, b\rangle$

Beweis: Wähle $c \in \mathbb{R}a + \mathbb{R}b$, damit ist $\det(a, b, c) = 0$, da a, b, c linear abhängig sind. Dann ist also $(a \times b, a) = (a \times b, b) = 0$ und daher a bzw. b senkrecht auf $a \times b$. Dies zeigt (i).

Zu (ii) Seien o.B.d.A a, b linear unabhängig. Wählen jetzt $c := \lambda(a \times b)$ mit $\lambda \in \mathbb{R}$ passend, so dass $||c|| = 1$ ist. Dann ist $(a \times b, c) = ||a \times b||$, andererseits $\mathrm{vol}\langle a, b, c\rangle = \mathrm{vol}\langle a, b\rangle$. $\square$

Bemerkung: Häufig ist folgende Beobachtung nützlich: Es ist $(a \times b) \times c = \lambda a + \mu b$, denn $(a \times b) \times c$ steht senkrecht auf $(a \times b)$ (und c), liegt daher in der Ebene $\mathbb{R}a + \mathbb{R}b$ (falls a, b linear unabhängig sind)