Lineare Algebra I

Prof. Dr. Dirk Ferus

Sommersemester 2001

Version vom 23.07.02

Inhaltsverzeichnis

Literatur

1. Fischer, G.: Lineare Algebra, vieweg
2. Köcher, M.: Lineare Algebra und analytische Geometrie, Springer (Enthält viele interessante historische Fakten)
3. Lorenz, F.: Lineare Algebra I, BI Hochschultaschenbücher
4. Jänich, K.: Lineare Algebra, Springer

1 Lineare Gleichungssysteme

Wir beginnen die Vorlesung mit dem Lösen von Gleichungen, genauer von mehreren Gleichungen für mehrere Unbekannte. Unter „Unbekannten“ verstehen wir dabei unbekannnte Zahlen, und Sie sollten sich einstweilen darunter reelle Zahlen (unendliche Dezimalbrüche) vorstellen. Später werden wir diesen Punkt präzisieren.

Zunächst zum Nachdenken folgendes

Beispiel 1. Wir betrachten das Gleichungssystem

$$\begin{array}{cccccc} x^4 & -3x^2 & +y^2 & +xy & +6 & =0 \\ & 4x^2 & -3y^2 & -3xy & -10 & =0 \end{array} .$$

Eine *Lösung* dieses Systems besteht aus *zwei* Zahlen x und y, die eingesetzt die *beiden* Gleichungen erfüllen. *Eine* Lösung ist also ein *geordnetes Paar* (x, y) von Zahlen.

Nun suchen wir eine Lösung.

Wir subtrahieren die zweite Zeile von der ersten, vertauschen die Seiten und erhalten

$$\begin{aligned} 0 &= x^4 - 7x^2 + 4y^2 + 4xy + 16 \\ &= x^4 - 8x^2 + 16 + x^2 + 4xy + 4y^2 \\ &= (x^2 - 4)^2 + (x + 2y)^2. \end{aligned}$$

Wir sehen daß $x^2 - 4 = 0$ und $x + 2y = 0$ sein müssen und erhalten

$$(x = 2, y = -1)$$

bzw.

$$(x = -2, y = 1).$$

Aber *keines* dieser beiden Paar löst auch nur *eine* der Ausgangsgleichungen! Der Grund dafür ist, daß wir im ersten Schritt unseres „Lösungsverfahrens“ Information verschenkt, d.h. Bedingungen ignoriert und damit die Lösungsmenge vergrößert haben. In Wirklichkeit hat das System gar keine reellen Lösungen, so wie etwa die Gleichung $x^2 + 1 = 0$ keine reellen Lösungen hat.

□

Die Frage, wie man die Lösungen eines Gleichungssystems findet, ist also offenbar nur *eine* Frage in diesem Zusammenhang. Andere wichtige Fragen sind,

- ob es überhaupt Lösungen gibt; andernfalls kann man sich jede Mühe sparen.
- ob es mehrere Lösungen gibt; in diesem Fall ist es interessant, zu wissen wieviele, und wie man sicher sein kann, alle gefunden zu haben.
- ob es vielleicht sogar unendlich viele Lösungen gibt, und wie man dann die Gesamtheit aller Lösungen beschreiben kann.

Wir untersuchen hier nicht beliebige Gleichungssysteme, sondern nur sogenannte *lineare Gleichungssysteme*. Eine *lineare Gleichung* ist eine Gleichung der Form

$$a_1x_1 \quad +a_2x_2 \quad +\dots \quad +a_nx_n \quad = b. \tag{1}$$

Dabei sind $a_{1,\dots}, a_n$ gegebene n Zahlen und $x_1, \dots, x_n$ gesuchte n Zahlen. Das Adjektiv *linear* bezieht sich darauf, daß nicht etwa Quadrate oder Produkte oder gar noch kompliziertere

Funktionen der Unbekannten in der Gleichung vorkommen. Hat man nur wenige Unbekannte, so bezeichnet man sie auch gern mit $x, y, z, \ldots$ statt mit $x_1, x_2, x_3, \ldots$.

Wir betrachten nun lineare Gleichungs*systeme*, also mehrere solche Gleichungen, wie etwa

$$\begin{array}{cccl} 3x & -2y & +5z & = 2 \\ 2x & +y & +9z & = 0. \end{array} \tag{2}$$

Definition 1. Ein *lineares Gleichungssystem* ist ein Gleichungssystem der Form

$$\begin{array}{ccccl} a_{11}x_1 & +a_{12}x_2 & +\ldots & +a_{1n}x_n & = b_1 \\ a_{21}x_1 & +a_{22}x_2 & +\ldots & +a_{2n}x_n & = b_2 \\ & & \ldots & & \\ a_{m1}x_1 & +a_{m2}x_2 & +\ldots & +a_{mn}x_n & = b_m \end{array} \tag{3}$$

Dabei sind die a_{ij} und die b_i als bekannt vorausgesetzt, und man hat dann m Gleichungen für die n Unbekannten $x_1, \ldots, x_n$. Eine *Lösung* dieses Systems ist ein geordnetes n-tupel $(x_1, \ldots, x_n)$ von Zahlen, das alle m Gleichungen erfüllt.

Für die linearen Gleichungssysteme gibt es eine sehr befriedigende Antwort auf alle obigen Fragen: Es gibt einen Algorithmus, ein allgemeines Verfahren, das

- zur Lösung führt, wenn es genau eine gibt,
- alle Lösungen liefert, falls es mehrere gibt,
- ggf. meldet, daß es keine Lösung gibt.

Dieses Verfahren ist der **Gaußsche Algorithmus**, den wir nun beschreiben. Wir führen ihn zunächst an einem Beispiel vor:

Beispiel 2. Wir betrachten

$$\begin{array}{cccl|l} 2x_1 & +x_2 & +3x_3 & = 0 & \\ 4x_1 & -x_2 & +x_3 & = 10 & -2 \times \text{erste Zeile} \\ 3x_1 & +2x_2 & +x_3 & = 7 & -\frac{3}{2} \times \text{erste Zeile} \end{array}$$

Wenn man wie angegeben verfährt, erhält man das neue System

$$\begin{array}{cccl|l} 2x_1 & +x_2 & +3x_3 & = 0 & \\ & -3x_2 & -5x_3 & = 10 & \\ & \frac{1}{2}x_2 & -\frac{7}{2}x_3 & = 7 & +\frac{1}{6} \times \text{zweite Zeile} \end{array}$$

Im nächsten Schritt erhält man

$$\begin{array}{cccl} 2x_1 & +x_2 & +3x_3 & = 0 \\ & -3x_2 & -5x_3 & = 10 \\ & & -\frac{13}{3}x_3 & = \frac{26}{3} \end{array} .$$

Beachten Sie, daß man bei diesen Umformungen keine Information verloren hat: Addiert man zum Beispiel zur 2. Zeile des zweiten Systems das Doppelte der ersten Zeile, so erhält man wieder die zweite Zeile des Ausgangssystems. Alle Umformungen sind umkehrbar. Deshalb ändert sich die Menge der Lösungen des Systems bei diesen Umformungen nicht, und das dritte System hat dieselbe Menge von Lösungen wie das erste. Aber das dritte System ist sehr leicht zu lösen: Aus der letzten Gleichung folgt notwendig $x_3 = -2$. Einsetzen in die vorletzte Gleichung liefert $x_2 = 0$, und Einsetzen der gefundenen Werte in die erste Gleichung des dritten Systems liefert $x_1 = 3$. Also hat unser System die *einzige* Lösung

$$(x_1, x_2, x_3) = (3, 0, -2).$$

□

Wir beschreiben nun den vorstehenden Algorithmus, den Gaußalgorithmus für lineare Gleichungssysteme, im allgemeinen.

Gegeben sei das System (3). Dabei wollen wir annehmen, daß x_1 wirklich in dem System vorkommt, daß also eines der a_{i1} von null verschieden ist. Sonst numerieren wir die Unbekannten um. Nach Vertauschung von zwei Zeilen können wir annehmen, daß sogar $a_{11} \neq 0$. Dann subtrahieren wir von den folgenden Zeilen ein Vielfaches der ersten, so daß jeweils x_1 verschwindet: genauer subtrahieren wir von der i-ten Zeile das $\frac{a_{i1}}{a_{11}}$-fache der ersten. Wir erhalten ein neues, zum ersten System äquivalentes, in dem x_1 nur noch in der ersten Zeile vorkommt.

Noch einmal die vorgenommenen **Umformungen**:

- Vertauschung von (zwei) Zeilen,
- Subtraktion des Vielfachen einer Zeile von einer anderen.

Beide Operationen sind umkehrbar: Durch nochmaliges Vertauschen bzw. durxh *Addition* desselben Vielfachen der einen Zeile zu der anderen erhält man das Ausgangssystem zurück. Bei Rechnungen von Hand ist es gelegentlich praktisch, noch folgende Umformung hinzuzunehmen:

- Multiplikation einer Zeile mit einem Faktor $\neq 0$.

Offenbar führt das auch zu einem äquivalenten System, und man kann eventuell Brüche vermeiden.

Bei den weiteren Umformungen des Systems betrachten wir nur noch die Zeilen ab der zweiten, die erste Zeile führen wir unverändert mit. Wir betrachten in den folgenden Zeilen die Unbekannte mit dem kleinsten noch vorkommenden Index. In der Regel wird das x_2 sein, es könnte aber auch unseren schon vorgenommenen Umformungen zum Opfer gefallen sein. Nehmen wir an, x_j wäre diese Unbekannte. Nach Zeilenvertauschung können wir wieder annehmen, daß x_j in der <u>zweiten</u> Zeile wirklich vorkommt. Dann subtrahieren wir wie eben von den Zeilen 3 bis m geeignete Vielfache der zweiten Zeile, so daß x_j <u>nur</u> noch in der zweiten Zeile vorkommt.

Dieses Verfahren setzen wir fort, bis in den nachfolgenden Zeilen gar keine Unbekannten mehr stehen. Das ist insbesondere dann der Fall, wenn es keine nachfolgenden Zeilen mehr gibt, sondern wir „unten angekommen sind". Es kann aber auch sein, daß noch mehr Gleichungen da sind, die aber auf der linken Seite keine Unbekannten mehr haben, also von der Form

$$0 = b \tag{4}$$

für irgendeine rechte Seite b sind. Unser System ist dann in Zeilen-Stufen-Form:

Definition 2. Ein lineares Gleichungssystem hat *Zeilenstufenform*, wenn folgendes gilt: Ist x_j die Unbekannte mit kleinstem Index in der i-ten Zeile, so kommt x_j in den folgenden Zeilen nicht mehr vor. Genauer:

$$\text{Aus } j = \min\{k \,|\, a_{ik} \neq 0\} \text{ folgt } a_{lj} = 0 \text{ für alle } l > i.$$

Ein System in Zeilenstufenform sieht so aus:

$$
\begin{array}{rl}
\boxed{\hspace{8em}} & = b_1 \\
\boxed{\hspace{5em}} & = b_2 \\
\boxed{\hspace{4.2em}} & = b_3 \\
\vdots & \\
\boxed{\hspace{2em}} & = b_r \\
0 & = b_{r+1} \\
\vdots & \\
0 & = b_m
\end{array}
$$

Wenn wir das ausschreiben wollen, gibt es ein Bezeichnungsproblem. Oben haben wir gesagt: „Ist x_j die Unbekannte mit kleinstem Index in der i-ten Zeile …“. Wenn wir nun aber über die ersten Unbekannnten in *mehreren* Zeilen gleichzeitig reden wollen, müssen wir „verschiedene Jots“ benutzen. Die mathematische Lösung dieses Bezeichnungsproblems bieten „gestufte“ Indizes. Wir bezeichnen die Unbekannte mit kleinstem Index in der i-ten Zeile nicht mit x_j, sondern mit x_{j_i}. Dabei ist j_i eben eine Zahl zwischen 1 und n.

Die i-te der ersten r Zeilen des oben schematisch dargestellten Systems in Zeilenstufenform sieht dann also so aus:

$$a_{ij_i}x_{j_i} + a_{ij_i+1}x_{j_i+1} + \ldots + a_{in}x_n = b_i, \qquad a_{ij_i} \neq 0,$$

wobei $j_1 < j_2 < \ldots < j_r$. Wir wollen x_{j_i} die Kopfvariable der Zeile i nennen. Beachten Sie: Ist dieses System das Ergebnis eines Gaußalgorithmus, so sind die a_{ij} und b_i im allgemeinen nicht mehr die Koeffizienten des Ausgangssystems. Wir müßten sie also eigentlich mit neuen Buchstaben wie $\tilde{a}_{ij}$ bezeichnen, was wir wegen der einfacheren Notation unterlassen. Im folgenden meinen wir also insbesondere mit b_i die rechte Seite der i-ten Zeile *in dem auf Zeilenstufenform transformierten Gleichungssystem.*

Die möglichen Fälle für die Lösungsmenge. Wegen der Umkehrbarkeit der vorgenommenen Manipulationen hat das neue System in Zeilenstufenform dieselben Lösungen wie das Ausgangssystem (3). Aber in der Zeilensttufenform kann man die Lösungsmenge ziemlich einfach ablesen:

<u>1. Fall:</u> $r < m$ und mindestens eine der Zahlen $b_{r+1}, \ldots, b_m$ ist nicht 0. Dann ist die entsprechende Zeile ein Widerspruch, und das Gleichungssystem hat

<u>keine Lösung</u>.

<u>2. Fall:</u> $b_{r+1} = \ldots = b_m = 0$, insbesondere $r = m$. Dann kann man die letzten $m - r$ Gleichungen der Form $0 = 0$ vergessen, sie enthalten keine Information. Das Gleichungssytem ist lösbar, und zwar wie folgt:

<u>Fall 2.A:</u> $r = n$. Dann sind alle Stufen von der Breite 1:

$$j_i = i.$$

Die letzte nichttriviale Gleichung lautet $a_{nn}x_n = b_n$ mit $a_{nn} \neq 0$, und wir erhalten daraus einen eindeutig bestimmten Wert x_n. Diesen setzen wir in die vorausgehenden Gleichungen ein. Aus

$$a_{n-1\,n-1}x_{n-1} + a_{n-1n}x_n = b_{n-1}, \quad a_{n-1\,n-1} \neq 0$$

ergibt sich dann x_{n-1} und so weiter. In diesem Fall hat das Gleichungssystem also

<u>eine eindeutig bestimmte Lösung</u> $(x_1, \ldots, x_n)$.

<u>Fall 2.B:</u> $r < n$. Dann gibt es wenigstens eine „breitere“ Stufe, und es gibt wenigstens eine Variable, die in keiner Zeile am Zeilenkopf steht. Solche Variablen nennen wir für den

Moment „freie Variable“. Diese freien Variablen kann man beliebig wählen und dann wieder von unten beginnend die Variablen an den Zeilenköpfen eindeutig bestimmen. In diesem Fall gibt es also

$$\underline{\text{unendlich viele Lösungen}}\ (x_1, \ldots, x_n).$$

Genauer kann man $n - r$ Variablen frei wählen, die anderen sind dann eindeutig bestimmt.

Wir fassen das noch einmal zusammen. Zunächst ist klar, daß

$$r \leq \min\{m, n\}.$$

Weiter gilt:

- Ist das System lösbar, so ist die Lösungsmenge durch $n - r$ frei wählbare „Parameter“ zu beschreiben. Dabei bedeutet $n - r = 0$ die eindeutige Lösbarkeit.
- Ist $r = m$ so gibt es (mindestens) eine Lösung, ist $r < m$, so kann das System lösbar oder unlösbar sein.

Beachten Sie aber: Der Gaußalgorithmus ist nicht eindeutig festgelegt. Zum Beispiel kann man sich bei den Zeilenvertauschungen einige Freiheit nehmen. Es ist nicht klar, ob die resultierende Zahl r von solcher Willkür unabhängig ist, oder ob man *ein* gegebenes System durch Gaußalgorithmen auf Zeilenstufenformen mit verschiedenem r bringen kann. Die vorstehenden Überlegungen zeigen allerdings, daß die Lösungsmenge des Ausgangssystems gewisse Bedingungen an r impliziert.

- Wenn das System keine Lösung hat, so ist $r < m$.
- Wenn das System genau eine Lösung hat, so isr $r = n$.

– egal, wie man den Gaußalgorithmus durchführt. Wir werden später sehen, daß r in Wahrheit nur vom Ausgangssystem abhängt. Es heißt der *Rang* des Gleichungssystems.

Wir betrachten nun einen Spezialfall linearer Gleichungssysteme.

Definition 3. 1. Das lineare Gleichungssystem (3) heißt *homogen*, wenn

$$b_1 = \ldots = b_m = 0.$$

2. Setzt man in einem beliebigen linearen Gleichungssystem (3) die rechten Seiten $= 0$, so erhält man eine neues lineares Gleichungssystem, das *zugehörige homogene* System.

Homogene lineare Gleichungssysteme haben immer eine Lösung, nämlich die sogenannte *triviale Lösung*

$$x_1 = \ldots = x_n = 0.$$

Wir beschließen diesen Abschnitt mit einem wichtigen Hilfssatz:

Lemma 1. *Hat das* homogene *lineare Gleichungssystem*

$$\begin{array}{lllll}
a_{11}x_1 & +a_{12}x_2 & +\ldots & +a_{1n}x_n & = 0 \\
a_{21}x_1 & +a_{22}x_2 & +\ldots & +a_{2n}x_n & = 0 \\
 & & \ldots & & \\
a_{m1}x_1 & +a_{m2}x_2 & +\ldots & +a_{mn}x_n & = 0
\end{array}$$

mehr Unbekannnte als Gleichungen, d.h. ist $m < n$, so besitzt das System eine nicht-triviale *Lösung*

$$(x_1, \ldots, x_n) \neq (0, \ldots, 0).$$

Beweis. Wir bringen das System mit dem Gaußalgorithmus auf Zeilenstufenform mit r nicht-trivialen Zeilen. Hätte es nur die triviale Lösung, also genau eine Lösung, so wäre nach unseren obigen Überlegungen $r = n$. Aber das steht im Widerspruch zu

$$r \leq m < n.$$

$\square$

2 Gruppen, Körper, Vektorräume

Bei den Überlegungen zum Gaußalgorithmus hätten wir statt reeller Zahlen ebenso gut rationale Zahlen oder komplexe Zahlen verwenden können. Die einzigen Rechenoperationen, die wir benutzt haben, sind Addition und Multiplikation mit ihren Umkehroperationen. Ganze Zahlen allein sind aber nicht ausreichend, eben weil man keine uneingeschränkte Division hat. So hat das lineare Mini-Gleichungssystem

$$2x = 3$$

keine Lösung *in den ganzen Zahlen*, obwohl das System bereits in Zeilenstufenform und $r = m = 1$ ist.

Wir werden nun genauer die Eigenschaften beschreiben, die wir von „Zahlen“ erwarten und benötigen. Wir betrachten Mengen mit „Rechenoperationen“, welche gewissen Axiomen genügen.

Definition 4 (Gruppe). Eine Menge G zusammen mit einer Abbildung (Verknüpfung)

$$G \times G \to G, (a, b) \mapsto ab$$

heißt eine *Gruppe*, wenn folgende Eigenschaften gelten:

1. $a(bc) = (ab)c$ für alle $a, b, c \in G$ (Assoziativgesetz).

2. Es gibt ein $e \in G$, so daß
   (a) $xe = ex = x$ für alle $x \in G$, und
   (b) zu jedem $x \in G$ ein $y \in G$ mit $xy = yx = e$ exisitiert.

Man kann zeigen, daß es in diesem Fall nur *ein* solches e gibt, das man *das neutrale Element* der Gruppe nennt, und daß es dann zu jedem $x \in G$ nur *ein* y mit der vorstehenden Eigenschaft gibt. Man nennt y *das Inverse* von x und schreibt dafür x^{-1}.

Als Beispiel für die Konsequenzen der Axiome beweisen wir, daß für alle $a, b \in G$ gilt

$$ab = a \implies b = e. \tag{5}$$

Aus $ab = a$ folgt nämlich

$$b = eb = (a^{-1}a)b = a^{-1}(ab) = a^{-1}a = e.$$

Beispiel 3. Die Menge $G = \mathbb{R}\backslash\{0\}$ der von null verschiedenen reellen Zahlen ist mit der üblichen Multiplikation eine Gruppe. Das neutrale Element ist $e = 1$ und das Inverse von x ist $\frac{1}{x}$.

$\square$

Beispiel 4. Die Menge $\mathbb{Z}$ der ganzen Zahlen ist mit der *Addition* als Verknüpfung eine Gruppe. Das neutrale Element ist $e = 0$ und das Inverse von x ist $-x$.

$\square$

Beispiel 5. Sei $M = \{1, 2, 3\}$ und $G := \{\sigma : M \to M | \sigma \text{ bijektiv}\}$ die Menge der eineindeutigen Abbildungen von M auf sich. So ein $\sigma \in G$ kann man charakterisieren, indem man einfach die Bilder $(\sigma(1), \sigma(2), \sigma(3))$ angibt. Die Verknüpfung

$$\sigma\tau := \sigma \circ \tau$$

sei die Hintereinanderausführung der beiden Abbildungen. Es ist also

$$(1, 3, 2)(2, 1, 3) = (3, 1, 2),$$
$$(2, 1, 3)(1, 3, 2) = (2, 3, 1).$$

Das neutrale Element ist die identische Abbildung $e = (1, 2, 3)$ und das Inverse ist eben die inverse Abbildung, z.B.

$$(2, 3, 1)^{-1} = (3, 1, 2).$$

Dieses Beispiel kann man auch mit beliebigen Mengen M statt $\{1, 2, 3\}$ betrachten, die Gruppe heißt die *Permutationsgruppe* der Menge M.

$\square$

Sie sehen, daß die Verknüpfung in einer Gruppe und die zugrundeliegende Menge von sehr verschiedener Natur sein können.

Definition 5 (Abelsche Gruppe). Eine Gruppe G heißt *kommutativ* oder *abelsch*, wenn für alle $x, y \in G$

$$xy = yx.$$

Die beiden ersten Beipiele oben sind abelsch, das dritte aber offenbar nicht.

Gruppen sind ein Begriff von fundamentaler Wichtigkeit an sehr vielen Stellen der Mathematik. In der linearen Algebra werden sie uns allerdings erst später wieder interessieren, wir haben es zunächst mit etwas komplizierteren Begriffen zu tun.

Definition 6 (Körper). Ein *Körper* ist eine Menge K zusammen mit *zwei* Verknüpfungsabbildungen

$$K \times K \to K, (x, y) \mapsto x + y,$$
$$K \times K \to K, (x, y) \mapsto xy,$$

die man *Addition* bzw. *Multiplikation* nennt, und die folgende Eigenschaften haben:

1. K mit der Addition ist eine abelsche Gruppe. Das neutrale Element bezeichnen wir mit 0 und das additive Inverse von x mit $-x$.

2. Für alle $x, y \in K \setminus \{0\} =: K_*$ ist auch $xy \in K_*$ und K_* ist mit der Multiplikation eine Gruppe. Das neutrale Element bezeichnen wir mit 1, das Inverse von x mit $x^{-1} = \frac{1}{x}$.

3. Für alle $x, y \in K$ ist $xy = yx$.

4. Addition und Multiplikation erfüllen das folgende sogenannte Distributivgesetz:

$$x(y + z) = xy + xz \text{ für alle } x, y, z \in K.$$

Beispiel 6. Die rationalen Zahlen $\mathbb{Q}$, die reellen Zahlen $\mathbb{R}$ und die komplexen Zahlen $\mathbb{C}$ bilden mit der üblichen Addition und Multiplikation jeweils einen Körper.

$\square$

Beispiel 7. Die Menge $\mathbb{Z}_2 = \{0, 1\}$ ist mit den Verknüpfungen

$$0 + 0 = 0, \quad 0 + 1 = 1 + 0 = 1, \quad 1 + 1 = 0,$$
$$0x = x0 = 0, \quad 1x = x1 = x \text{ für alle } x \in \{0, 1\}$$

ein Körper.

$\square$

Bemerkung: Im vorstehenden Beispiel ist also $1 + 1 = 0$, wobei 1 das neutrale Element der Multiplikation und 0 das neutrale Element der Addition in K bezeichnet. Allgemeiner ist in einem Körper K die k-fache Summe $1 + \ldots + 1$ entweder für alle $k \in \mathbb{N}$ von 0 verschieden - wie etwa bei den rationalen Zahlen -, oder es gibt eine Primzahl p, so daß die k-fache Summe $1 + \ldots + 1 = 0$, wenn immer k ein Vielfaches von p ist. Im ersten Fall sagt man, K ist ein *Körper der Charakteristik* 0, andernfalls einer der *Charakteristik p*. Die letzteren sind wichtig in der Zahlentheorie. Wir werden es aber vor allem mit Körpern der Charakteristik 0 zu tun haben. In einem Körper hat man also die Grundrechenarten mit den geläufigen Rechenregeln zur Verfügung. Und mit einem beliebigen Körper anstelle von $\mathbb{R}$ lassen sich zum Beispiel der Gaußsche Algorithmus und die Theorie linearer Gleichungssysteme wie oben formulieren.

Die axiomatische Methode. Die vorstehenden Definitionen der Begriffe „Gruppe“ und „Körper“ sind Beipiele für die sogenannte axiomatische Methode, mathematische Begriffe einzuführen. Auf der einen Seite ist das ein sehr abstraktes Verfahren, auch im historischen Sinne: die Axiome für eine Gruppe oder einen Körper hat man sich nicht einfach ausgedacht. Niemand hat gesagt: „Schreiben wir mal ein paar Axiome hin und sehen was daraus folgt.“ Vielmehr ist der Entscheidung für ein solches System von Axiomen eine lange (ohne Übertreibung: jahrhundertelange) Erfahrung im Umgang zum Beipiel mit Gleichungen verschiedener Art vorausgegangen, begleitet vom Abstraktionsprozeß: „Worauf kommt es wirklich an?“

Das Ergebnis kann man aber auch sehr pragmatisch sehen: Statt uns auf die diffizile philosophische Frage einzulassen, was Zahlen eigentlich sind, formulieren wir die Rechenregeln, die wir für sie voraussetzen wollen: Eben die Axiome für einen Körper. Die Frage, ob es wirklich „Zahlen“ gibt, die den Regeln für einen Körper genügen, lassen wir im Rahmen dieser Vorlesung dahingestellt sein. Wir akzeptieren die Existenz zum Beipiel der rationalen oder reellen Zahlen. Wir werden im folgenden sehr viel mit Körpern zu tun haben. Einstweilen verpassen Sie aber nichts, wenn Sie sich dann einfach die reellen Zahlen vorstellen.

Die wichtigsten Objekte der linearen Algebra sind Vektorräume über einem Körper. Für „Vektoren“ gilt das eben über Zahlen Gesagte genauso. Definitionen wie „Vektoren sind gerichtete Strecken im Raum“ oder „Größen, die eine Länge und eine Richtung haben“ sind problematisch: Was ist denn eine „Richtung“? Wir stellen uns wieder auf den pragmatischen Standpunkt: Wir erklären zunächst einmal, wie man mit Vektoren rechnet, indem wir axiomatisch *Vektorräume* definieren. Und dann diskutieren wir verschiedene Modelle, in denen diese Rechenregeln gelten, Modelle für Vektorräume. Diese Modelle geben ganz verschiedene Antworten auf die Frage, wie man sich Vektoren vorstellen soll (oder kann), und sie stehen für sehr verschiedene Anwendungsbereiche des Begriffs.

Definition 7 (Vektorraum). Sei K ein Körper. Ein *Vektorraum über* K oder ein *K-Vektorraum* ist eine Menge V zusammen mit zwei Abbildungen

$$V \times V \to V, (v, w) \mapsto v + w,$$
$$K \times V \to V, (\lambda, v) \mapsto \lambda v,$$

die man *(Vektor-)Addition* und *Skalarmultiplikation* nennt, so daß gilt:

1. V ist zusammen mit der Addition eine abelsche Gruppe. Das neutrale Element bezeichnen wir mit 0 oder, um es vom Nullelement des Körpers K zu unterscheiden, mit 0_V. Wir nennen es auch *den Nullvektor* von V.

2. Mit beliebigen $v, w \in V$ und $\lambda, \mu \in K$ ist

$$\lambda(\mu v) = (\lambda\mu)v,$$
$$\lambda(v + w) = \lambda v + \lambda w, \quad (\lambda + \mu)v = \lambda v + \mu v,$$
$$1v = v.$$

   Dabei ist 1 das neutrale Element der Multiplikation in K. Die Elemente von K bezeichnet man häufig auch als *Skalare*, die Elemente von V als *Vektoren.*

Jetzt können wir die Frage beantworten, was ein Vektor ist: ein Element eines Vektorraumes.

Als Beispiel für die Anwendung der Axiome beweisen wir, daß für alle $v \in V$

$$0v = 0_V$$
$$(-1)v = -v.$$

Es ist nämlich

$$0v = (0 + 0)v = 0v + 0v.$$

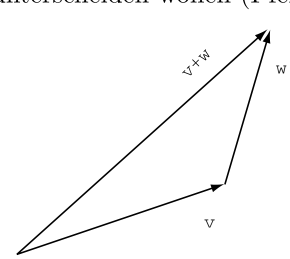
Aus (5) angewendet auf die abelsche Gruppe V mit der Addition folgt $0v = 0$, und die erste Gleichung ist bewiesen. weiter ist

$$v + (-1)v = 1v + (-1)v = (1 + (-1))v = 0v = 0_V,$$

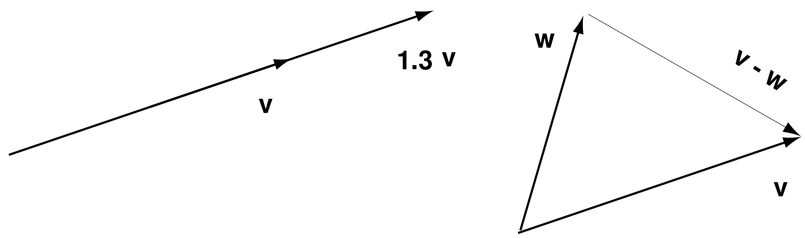
und aus der Eindeutigkeit des (additiven) Inversen folgt $(-1)v = -v$.

Beispiel 8 (Ein geometrisches Modell). Dieses Modell begegnet Ihnen fast überall, wo es um geometrische Veranschaulichungen geht. Zum Beispiel bei der Beschreibung von Kräften an einer mechanischen Konstruktion (Kräfte-Diagramm), bei der Bewegung von Teilchen im Raum oder der Beschreibung von Kraftfeldern. Entsprechend appelliert dieses Modell an die Anschauung und ist mathematisch nicht so ganz exakt definiert. Der Körper sind die reellen Zahlen $\mathbb{R}$. Wir betrachten Pfeile in der Ebene (oder im Raum), wobei wir zwischen parallelverschobenen Pfeilen nicht unterscheiden wollen (Pfeilklassen).

Die Addition erklären wir so:



Die Multiplikation mit $\lambda \in \mathbb{R}$ ist die Streckung des Pfeils um den Faktor a, wobei negatives Vorzeichen die Pfeilrichtung umkehrt. Man hat also:



Zur Übung sollten Sie sich die Axiome anschaulich klar machen. Was bedeutet zum Beispiel das Gruppenaxiom $u + (v + w) = (u + v) + w$?

□

Beispiel 9 (Ein algebraisches Modell). Dieses Modell ist *das* Standardmodell für einen Vektorraum. Sei K ein Körper und $n \in \mathbb{N}$ eine natürliche Zahl. Wir betrachten den K^n, das ist die Menge geordneter „n-tupel“ von Elementen aus K:

$$K^n = \{x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \mid x_i \in K\}.$$

Für $n = 0$ vereinbaren wir, daß K^0 die Menge mit dem einzigen Element $0 \in K$ sein soll. Die Addition und Multiplikation mit (reellen) Skalaren sind komponentenweise erklärt:

$$\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} + \begin{pmatrix} y_1 \\ \vdots \\ {}_n \end{pmatrix} = \begin{pmatrix} x_1 + y_1 \\ \vdots \\ x_n + y_n \end{pmatrix}$$

$$\lambda \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} \lambda x_1 \\ \vdots \\ \lambda x_n \end{pmatrix}$$

Die oben geforderten Rechenregeln ergeben sich dann einfach aus den entsprechenden Regeln in K.

Bemerkung zum Fall $n = 1$: Vektoren im K^1 sind von der Form (x), wobei $x \in K$. Läßt man die Klammer weg, so sind die Addition von Vektoren und die Skalarmultiplikation einfach die übliche Addition und Multiplikation. Der Körper K ist also auf natürliche Weise auch einen K-Vektorraum.

□

Beispiel 10 (Polynome). Wir betrachten den Körper $K = \mathbb{R}$ und bezeichnen mit V die Menge der Polynome auf $\mathbb{R}$, d.h. die Menge der Funktionen $\mathbb{R} \to \mathbb{R}$, die sich schreiben lassen in der Form

$$f : x \mapsto a_0 + a_1 x + a_2 x^2 + \ldots + a_n x^n$$

für ein $n \in \mathbb{N}$ und *Koeffizienten* $a_0, \ldots, a_n \in \mathbb{R}$. Summen und skalare Vielfache solcher Funktionen sind auf die offensichtliche Weise (nämlich wertweise) erklärt, und die Polynome bilden einen Vektorraum über $\mathbb{R}$. Statt $\mathbb{R}$ kann man auch einen beliebigen Körper nehmen. (Anmerkung für Kenner: Für Körper der Charakteristik $\neq 0$ ist der so definierte Funktionenraum *nicht* der in der Algebra übliche Raum der Polynome.)

□

Das vorstehende Beispiel läßt sich verallgemeinern:

Beispiel 11 (Funktionen). Sei K ein Körper, M eine nichtleere Menge und

$$V := \{f : M \to K | f \text{ eine beliebige Abbildung}\}.$$

Dann ist V mit der wertweise definierten Addition und Skalarmultiplikation ein K-Vektorraum, der Vektorraum der K-wertigen Funktionen auf M. Man schreibt auch $V = K^M$. Beachten Sie, daß der K^n „dasselbe" ist wie $K^{\{1,\dots,n\}}$.

□

Definition 8 (Untervektorraum). Sei V ein K-Vektorraum und $U \subset V$. Dann heißt U ein *Untervektorraum von* V, wenn gilt: Für alle $x, y \in U$ und $\lambda \in K$ sind

$$x + y \in U \text{ und } \lambda x \in U,$$

und

$$U \neq \emptyset.$$

<u>Bemerkung:</u> Weil $U \neq \emptyset$, gibt es in diesem Fall ein $x \in U$. Nach Voraussetzung ist dann

$$(-1)x = -x \in U$$

und daher auch

$$0_V = x + (-x) \in U.$$

Daher kann man die Rechenoperationen von V einfach auf U einschränken und U wird mit diesen Operationen selbst ein K-Vektorraum. Daher der Name *Untervektorraum*.

Beispiel 12. Der Vektorraum K^M der K-wertigen Funktionen auf der Menge M enthält den Untervektorraum

$$K^{[M]} := \{f : M \to K \mid f(x) \neq 0 \text{ nur für endlich viele } x \in M\}.$$

□

Beispiel 13 (Homogene lineare Gleichungssysteme). Sei K ein Körper. Wir betrachten das homogene lineare Gleichungssystem

$$\begin{array}{llll} a_{11}x_1 & +a_{12}x_2 & +\dots & +a_{1n}x_n & = 0 \\ a_{21}x_1 & +a_{22}x_2 & +\dots & +a_{2n}x_n & = 0 \\ & & \dots & & \\ a_{m1}x_1 & +a_{m2}x_2 & +\dots & +a_{mn}x_n & = 0 \end{array} \tag{6}$$

mit $a_{ij} \in K$ und gesuchten Lösungen $\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in K^n$. Wegen der Rechenregeln in K sind mit zwei Lösungen $\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$ und $\lambda \in K$ auch

$$\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} + \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} x_1 + y_1 \\ \vdots \\ x_n + y_n \end{pmatrix}$$

und

$$\lambda \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} \lambda x_1 \\ \vdots \\ \lambda x_n \end{pmatrix}$$

zwei Lösungen des Systems. Weiter ist die triviale Lösung $(0, \ldots, 0)$ eine Lösung, die Lösungsmenge also nicht leer. Damit gilt der

Satz 1. *Die Lösungsmenge eines homogenen linearen Gleichungssystems über K mit $n > 0$ Unbekannten ist ein Untervektorraum des K^n.*

$\square$

Beispiel 14 (Homogene lineare Differentialgleichungen). Sei $V = \mathbb{R}^{\mathbb{R}}$ der $\mathbb{R}$-Vektorraum der reellen Funktionen auf $\mathbb{R}$ und sei

$$U := \{y \in V \mid y \text{ ist } 2\times \text{ differenzierbar und } y'' + y = 0\}.$$

Zum Beispiel ist $y = \sin$ ein Element von U. Wie man leicht sieht, ist U ein Untervektorraum von V. Allgemeiner ist der Lösungsraum einer homogenen linearen Differentialgleichung auf $\mathbb{R}$ eine Untervektorraum von V.

$\square$

Ein wesentliches Ziel für uns wird sein, die Beziehungen zwischen den beiden letzten Beispielen zu verstehen.

Zwei Methoden, aus Vektorunterräumen neue Vektorunterräume zu basteln:

Definition 9. Seien U_1, U_2 zwei Untervektorräume des K-Vektorraumes V. Dann sind auch

$$U_1 \cap U_2$$

und

$$U_1 + U_2 := \{u_1 + u_2 \mid u_1 \in U_1 \text{ und } u_2 \in U_2\}$$

Vektorunterräume von V. Man bezeichnet sie als den *Durchschnitt* und die *Summe* von U_1 und U_2.

Ist

$$U_1 \cap U_2 = \{0\},$$

so schreibt man statt $U_1 + U_2$ auch $U_1 \bigoplus U_2$ und nennt diesen Untervektrorraum *die direkte Summe* von U_1 und U_2. In diesem Fall ist die Darstellung

$$u = u_1 + u_2, \quad u_i \in U_i$$

für $u \in U_1 \bigoplus U_2$ eindeutig.

Beweis der Eindeutigkeit. Aus

$$u_1 + u_2 = v_1 + v_2, \quad u_i, v_i \in U_i$$

folgt

$$u_1 - v_1 = v_2 - u_2.$$

Die linke Seite ist in U_1, die rechte in U_2, und weil $U_1 \cap U_2 = \{0\}$, ist

$$u_1 - v_1 = 0 = v_2 - u_2,$$

also $u_1 = v_1$ und $u_2 = v_2$. $\square$

3 Linearkombinationen

Definition 10 (Linearkombinationen). Seien V ein K-Vektorraum, $k \in \mathbb{N}$, $v_1, \ldots, v_k \in V$ Vektoren und $\lambda_1, \ldots, \lambda_k \in K$ Skalare. Dann kann man Skalarmultiplikation und Addition mehrfach ausführen und den Vektor

$$v = \lambda_1 v_1 + \ldots + \lambda_k v_k$$

bilden. Jedes $v \in V$, das sich auf diese Weise schreiben läßt, heißt eine *Linearkombination* der Vektoren $v_1, \ldots, v_k$.

Beispiel 15. Im $\mathbb{R}^3$ betrachten wir die drei Vektoren

$$x = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \; y = \begin{pmatrix} -1 \\ 2 \\ 1 \end{pmatrix}, \; z = \begin{pmatrix} 0 \\ 2 \\ 2 \end{pmatrix}.$$

Dann ist $0 \in \mathbb{R}^3$ eine Linearkombination von x, y, z, nämlich

$$0 = 0x + 0y + 0z.$$

Das hat mit x, y, z herzlich wenig zu tun und heißt auch *die triviale Linearkombination*. Aber es ist auch

$$0 = 3x + 3y - 3z.$$

Sie sehen: Die Darstellung eines Vektors als Linearkombination gegebener Vektoren $v_1, \ldots, v_k$ ist nicht unbedingt eindeutig.

□

Beispiel 16. Im K^n betrachten wir die Vektoren

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

Der Vektor e_i hat also an der i-ten Stelle eine 1 und sonst lauter Nullen. Dann ist nach der Definition der Rechenoperationen im K^n für jedes $x = (x_1, \ldots, x_n)$

$$x = x_1 e_1 + \ldots + x_n e_n.$$

Jeder Vektor des K^n ist also eine Linearkombination von $e_1, \ldots, e_n$. In diesem Fall ist die Darstellung auch eindeutig: Nur e_1 kann zur ersten Komponente x_1 von x beitragen, und darum ist der Koeffizient von e_1 notwendig x_1. Entsprechendes gilt für die anderen Komponenten.

□

Beispiel 17. Die Funktion

$$x^3 \in \mathbb{R}^{\mathbb{R}}$$

ist eine Linearkombination der Funktionen

$$(x-1)^k, \; k = 0, 1, 2, 3.$$

Es ist nämlich (Nachrechnen!)

$$x^3 = 1\underbrace{(x-1)^0}_{v_1} + 3\underbrace{(x-1)^1}_{v_2} + 3\underbrace{(x-1)^2}_{v_3} + 1\underbrace{(x-1)^3}_{v_4}.$$

Umgekehrt und allgemeiner ist die Funktion $(1+x)^n$ eine Linearkombination der Monome $1, x, x^2, \ldots, x^n$. Die Koeffizienten sind die sogenannten Binomialkoeffizienten $\binom{n}{k}$:

$$(1+x)^n = \binom{n}{0}1 + \binom{n}{1}x + \binom{n}{2}x^2 + \ldots + \binom{n}{n}x^n = \sum_{k=0}^{n}\binom{n}{k}x^k.$$

□

Beispiel 18. Die Funktion $\sin(t+\frac{\pi}{3})$ ist eine Linearkombination der Funktionen $\sin t$ und $\cos t$:

$$\sin(t+\frac{\pi}{3}) = \sin t \cos\frac{\pi}{3} + \cos t \sin\frac{\pi}{3} = \frac{1}{2}\sin t + \frac{\sqrt{3}}{2}\cos t.$$

□

Definition 11. Ist $M \subset V$ eine beliebige Teilmenge, so sagt man $v \in V$ ist eine Linearkombination von Vektoren (oder Elementen) aus M, wenn es in M *endlich viele* Elemente $v_1, \ldots, v_k$ gibt, so daß v eine Linearkombination von diesen ist. In diesem Fall kann man überdies annehmen, daß $v_1, \ldots, v_k$ paarweise verschieden sind, weil man die Koeffizienten eines mehrmals auftretenden Vektors ja zusammenfasssen kann.

Bemerkungen. Es ist hilfreich, den Nullvektor 0_V als Linearkombination von Elementen der leeren Menge $M = \emptyset$ zu definieren. (Leere Summe= 0.)

Eine Teilmenge $U \subset V$ ist offenbar genau dann ein Untervektorraum von V, wenn jede Linearkombination von Elementen aus U wieder in U liegt, d.h. wenn U *abgeschlossen gegenüber Linearkombinationen* ist.

Definition 12 (Lineare Hülle, Spann). Seien V ein K-Vektorraum und $M \subset V$ eine Teilmenge. Sei

$$\mathrm{Spann}(M) := \bigcap_{M \subset U} U,$$

wobei der Durchschnitt gebildet ist über alle Untervektorräume $U \subset V$, die M enthalten. $\mathrm{Spann}(M)$ heißt *die lineare Hülle* oder*der Spann von M*. Für $v_1, \ldots, v_k \in V$ schreibt man

$$\mathrm{Spann}(v_1, \ldots, v_k) := \mathrm{Spann}(\{v_1, \ldots, v_k\}).$$

Beispiel 19.

$$\mathrm{Spann}(\emptyset) = \{0\}.$$

□

Satz 2. *Mit den Voraussetzungen und Bezeichnungen der Definition ist* $\mathrm{Spann}(M)$ *ein Vektorunterraum von V und es gilt:*

$$\mathrm{Spann}(M) = \{v \in V \mid v \textit{ ist Linearkombination von Vektoren aus } M\}.$$

Beweis. Sind $v_1, v_2, v \in \text{Spann}(M)$, so liegen sie in jedem Untervektorraum, der M enthält. Also liegen $v_1 + v_2$ und für jedes $\lambda \in K$ auch λv in jedem Unterraum, der M enthält und damit in $\text{Spann}(M)$. Also ist $\text{Spann}(M)$ ein Untervektorraum.

Wir setzen

$$U_M := \{v \in V \mid v \text{ ist Linearkombination von Vektoren aus } M\}.$$

Ist $v \in U_M$ eine Linearkombination der Vektoren $v_1, \ldots, v_k \in M$, so liegen diese Vektoren natürlich in *jedem* Untervektorraum von V, der M enthält. Damit liegt auch v in *jedem* Untervektorraum, der M enthält. Daher ist

$$U_M \subset \text{Spann}(M) \tag{7}$$

$$\tag{8}$$

Andrerseits ist U_M ein Untervektorraum, weil die Summe von zwei Linearkombinationen von Vektoren aus M und skalare Vielfache einer solchen Linearkombination wieder Linearkombinationen von Vektoren aus M sind. Da offenbar $M \subset U_M$ folgt

$$\text{Spann}(M) \subset U_M. \tag{9}$$

Aus (7) und (9) folgt

$$\text{Spann}(M) = U_M.$$

Weil wir schon gezeigt haben, daß U_M ein Untervektorraum von V ist, gilt das dann auch für $\text{Spann}(M)$. □

Beispiel 20. Als Konsequenz aus diesem Satz und Beispiel 16 ergibt sich

$$K^n = \text{Spann}(e_1, \ldots, e_n).$$

□

Die Aufgabe, einen Untervektorraum durch möglichst wenige Vektoren zu erzeugen, führt zum fundamentalen Begriff der *linearen Unabhängigkeit.*

Wir treffen folgende Konvention: Ein „Dach“ über einem Glied eines k-tupels soll bedeuten, daß dieser Term ausgelassen wird:

$$\boxed{(v_1, \ldots, \widehat{v_i}, \ldots, v_k) := (v_1, \ldots, v_{i-1}, v_{i+1}, \ldots, v_k).}$$

Definition 13. Seien V ein K-Vektorraum, $M \subset V$ und $v_1, \ldots, v_k \in V$.

1. M heißt *linear unabhängig*, wenn für alle $u \in M$
$$\text{Spann}(M \backslash \{u\}) \neq \text{Spann}(M).$$
2. Das k-tupel $(v_1, \ldots, v_k)$ oder die Vektoren $v_1, \ldots, v_k$ heißen *linear unabhängig*, wenn für alle $i \in \{1, \ldots, k\}$
$$\text{Spann}(v_1, \ldots, \widehat{v_i}, \ldots, v_k) \neq \text{Spann}(v_1, \ldots, v_k).$$

Sind M oder $v_1, \ldots, v_k$ nicht linear unabhängig, so nennt man sie *linear abhängig.*

Beispiel 21. Die Vektoren $e_1, \ldots, e_n \in K^n$ sind linear unabhängig. Alle Linearkombinationen der $e_1, \ldots, e_n$, die e_i nicht enthalten, haben nämlich an der i-ten Stelle eine 0. Deshalb ist

$$\text{Spann}(\{e_1, \ldots, e_n\} \backslash \{e_i\}) \neq K^n = \text{Spann}(\{e_1, \ldots, e_n\}).$$

□

Beispiel 22. Ein einzelner Vektor v ist linear unabhängig genau dann, wenn

$$\mathrm{Spann}(v) = \{\lambda v \mid \lambda \in K\} \neq \mathrm{Spann}(\emptyset) = \{0\}$$

ist, d.h. wenn $v \neq 0$ ist.

□

Satz 3 (Kriterien für lineare Unabhängigkeit). *Seien V ein K-Vektorraum, $M \subset V$ und $v_1, \ldots, v_k \in V$.*

1. *M ist linear unabhängig genau dann, wenn kein $u \in M$ Linearkombination von* ande*ren Vektoren aus M ist.*

2. *M ist linear unabhängig, wenn 0_V sich nur auf die triviale Weise als Linearkombination von Vektoren aus M schreiben läßt, d.h. wenn für alle paarweise verschiedenen $u_1, \ldots, u_k \in M$*

$$\lambda_1 u_1 + \ldots + \lambda_k u_k = 0_V \implies \lambda_1 = \ldots = \lambda_k = 0.$$

3. *$(v_1, \ldots, v_k)$ ist linear unabhängig genau dann, wenn keiner dieser Vektoren eine Linearkombination der anderen ist.*

4. *$(v_1, \ldots, v_k)$ ist linear unabhängig genau dann, wenn sich 0_V nur trivial aus $v_1, \ldots, v_k$ linear kombinieren läßt, d.h. wenn*

$$\lambda_1 v_1 + \ldots + \lambda_k v_k = 0_V \implies \lambda_1 = \ldots = \lambda_k = 0$$

gilt.

Beweis. Zu 1. $u \in M$ ist Linearkombination von *anderen* Vektoren aus M genau dann, wenn

$$u \in \mathrm{Spann}(M \backslash \{u\}.$$

Weil $u \in M$ ist dies äquivalent zu

$$\mathrm{Spann}(M) = \mathrm{Spann}(M \backslash \{u\}.$$

Wir haben damit gezeigt: Es gibt ein $u \in M$, welches Linearkombination von *anderen* Vektoren aus M ist, genau dann, wenn M linear abhängig ist. Das ist äquivalent zur Behauptung.

Zu 2. Ist M linear abhängig, so gibt es nach 1. ein $u \in M$ und paarweise verschiedene Vektoren $u_1, \ldots, u_k \in M \backslash \{u\}$ und weiter $\lambda_1, \ldots, \lambda_k \in K$, so daß

$$u = \lambda_1 u_1 + \ldots \lambda_k u_k,$$

also

$$0_V = (-1)u + \lambda_1 u_1 + \ldots \lambda_k u_k.$$

Gibt es andrerseits eine Linearkombination

$$0 = \lambda_1 u_1 + \ldots + \lambda_k u_k$$

mit paarweise verschiedenen $u_1, \ldots, u_k \in M$ und $\lambda_i \neq 0$, so ist

$$u_i = -\sum_{j \neq i} \frac{\lambda_j}{\lambda_i} u_j$$

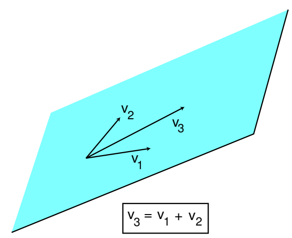
eine Linearkombination von Vektoren aus $M\backslash\{u_i\}$, also M linear abhängig. Wir haben damit gezeigt: M ist genau dann linear abhängig, wenn 0 eine nicht-triviale Linearkombination von Vektoren aus M ist. Daraus folgt 2.

Zu 3. Folgt direkt aus 1.

Zu 4. Folgt direkt aus 2. $\square$

Beispiel 23. Zwei Vektoren $v_1, v_2 \in V$ sind linear unabhängig, wenn keiner ein Vielfaches des anderen ist. Geometrisch bedeutet das etwa im $\mathbb{R}^3$, daß die beiden Vektoren nicht auf einer Geraden durch 0 liegen.

Drei Vektoren im $\mathbb{R}^3$ sind linear abhängig, wenn sie in einer Ebene durch 0 liegen.



$\square$

Beispiel 24. Vgl. Beispiel 21. Die Vektoren $e_1, \ldots, e_n$ des K^n sind linear unabhängig. Wäre nämlich e_i eine Linearkombination der anderen, so wäre die i-Komponente von e_i nämlich 0. Sie ist aber 1.

$\square$

Beispiel 25. Sei $n \in \mathbb{N}$. Die Funktionen (Monome)

$$1, x, x^2, \ldots, x^n \in \mathbb{R}^{\mathbb{R}}$$

sind linear unabhängig. Wir beweisen das durch vollständige Induktion über n.

$\underline{n = 0.}$ Die Funktion 1 ist nicht 0, also linear unabhängig.

$\underline{n \to (n+1).}$ Die Behauptung sei für n bewiesen und es sei

$$\lambda_0 1 + \lambda_1 x + \ldots + \lambda_{n+1} x^{n+1} = 0.$$

(Auf der rechten Seite steht die Nullfunktion $0 : \mathbb{R} \to \mathbb{R}$.) Zu zeigen ist

$$\lambda_0 = \ldots = \lambda_{n+1} = 0.$$

Wir setzen $x = 0$ und finden

$$\lambda_0 = 0.$$

Also ist

$$x(\lambda_1 + \ldots + \lambda_{n+1} x^n) = 0$$

für alle x. Nach Division durch x findet man für alle $x \neq 0$

$$\lambda_1 + \lambda_2 x + \ldots + \lambda_{n+1} x^n = 0,$$

und weil beide Seiten stetig sind, gilt das auch für $x = 0$. Nach Induktionsvoraussetzung folgt

$$\lambda_1 = \ldots = \lambda_{n+1} = 0.$$

$\square$

Beispiel 26 (Unabhängigkeitstest im K^n). Seien

$$x_1 = \begin{pmatrix} x_{11} \\ \vdots \\ x_{n1} \end{pmatrix}, \ldots, x_k = \begin{pmatrix} x_{1k} \\ \vdots \\ x_{nk} \end{pmatrix}$$

k Vektoren im K^n. Um diese auf lineare Unabhängigkeit zu testen betrachten wir

$$\lambda_1 x_1 + \ldots + \lambda_k x_k = 0$$

als lineares Gleichungssystem für die λ_i:

$$\begin{array}{ccccc} x_{11}\lambda_1 & +x_{21}\lambda_2 & +\ldots & +x_{1k}\lambda_k & = 0 \\ & & \ldots & & \\ x_{n1}\lambda_1 & +x_{2n}\lambda_2 & +\ldots & +x_{nk}\lambda_k & = 0. \end{array}$$

Die Vektoren sind linear unabhängig genau dann, wenn dieses nur die Lösung $\lambda_1 = \ldots = \lambda_k = 0$ besitzt, d.h. wenn der Rang $r = k$ ist. Wendet man den Gaußalgorithmus an, so müssen also k nicht-triviale Zeilen übrigbleiben. Genau dann sind die Vektoren linear unabhängig. Zum Beispiel sind mehr als n Vektoren im K^n stets linear abhängig.

□

Wir kommen zurück zum Problem, alle Vektoren eines gegebenen Vektor(unter)raums als Linearkombinationen aus möglichst wenigen vorgegebenen Vektoren darzustellen. Dafür müssen wir den Begriff der linearen Unabhängigkeit noch etwas ergänzen.

Definition 14 (Erzeugendensystem, Basis). Sei M eine Teilmenge des Vektorraums V.

1. M heißt *ein Erzeugendensystem* von V, wenn
$$\text{Spann}(M) = V.$$
2. M heißt eine Basis von V, wenn M linear unabhängig und ein Erzeugendensystem von V ist.
3. Entsprechend nennt man ein k-tupel $(v_1, \ldots, v_k)$ ein Erzeugendensystem bzw. eine Basis von V.
4. V heißt *endlich erzeugt*, wenn es eine *endliche* Menge M gibt, die V erzeugt. V heißt von $n \in \mathbb{N}$-Elementen erzeugt, wenn es ein Erzeugendensystem von V mit $\leq n$ Elementen gibt. (Eventuell reichen also auch weniger als n Elemente aus.)

Beispiel 27 (Die Standardbasis des K^n). Die Vektoren

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

bzw. das n-tupel $(e_1, \ldots, e_n)$ des K^n sind nach Beispiel 20 ein Erzeugendensystem und nach Beispiel 21 bzw. 24 linear unabhängig. Man nennt sie *die Standardbasis* des K^n.

□

Beispiel 28. Die leere Menge ist eine Basis des Nullvektorraums $V = \{0\}$. In diesem Zusammenhang wollen wir einen Ausdruck $(b_1, \ldots, b_r)$ mit $\underline{r = 0}$ als leere Menge interpretieren.

$\square$

Beispiel 29 (Monom-Basis). Die Menge der Monome $M := \{1, x, x^2, \ldots\}$ ist linear unabhängig. (Wie folgt das aus Beispiel 25?) Die Linearkombinationen aus endlich-vielen dieser Funktionen sind gerade die Polynome. Also ist die Menge M der Monome eine Basis des Raums

$$\mathrm{Pol}(\mathbb{R}) := \mathrm{Spann}(M)$$

der Polynome auf $\mathbb{R}$.

$\square$

Lemma 2. *Sind $(b_1, \ldots, b_r)$ linear unabhängig in V und ist*

$$b_{r+1} \in V \setminus \mathrm{Spann}(b_1, \ldots, b_r),$$

so sind $(b_1, \ldots, b_r, b_{r+1})$ linear unabhängig.

Beweis. Sei

$$\lambda_1 b_1 + \ldots + \lambda_r b_r + \lambda_{r+1} b_{r+1} = 0.$$

Wir wollen zeigen, daß alle $\lambda_i = 0$. Zunächst ist $\lambda_{r+1} = 0$, denn sonst könnte man die Gleichung nach b_{r+1} auflösen, und es wäre $b_{r+1} \in \mathrm{Spann}(b_1, \ldots, b_r)$ im Widerspruch zur Voraussetzung. Damit folgt

$$\lambda_1 b_1 + \ldots + \lambda_r b_r = 0,$$

und wegen der linearen Unabhängigkeit von $(b_1, \ldots, b_r)$ sind auch $\lambda_1 = \ldots = \lambda_r = 0$. $\square$

Lemma 3 (Schrankenlemma). *Seien V ein K-Vektorraum und $n \in \mathbb{N}$. Dann sind folgende Aussagen äquivalent:*

1. *V wird von n Elementen erzeugt.*
2. *Je $n + 1$ Elemente von V sind linear abhängig.*

Beweis. Ist $V = \{0\}$, so ist nichts zu zeigen. Wir nehmen also an, daß $V \neq \{0\}$.

$\underline{1 \implies 2.}$ Seien $v_1, \ldots, v_{n+1} \in V$. Wir wollen zeigen, daß es $\lambda_1, \ldots, \lambda_{n+1} \in K$ gibt, die nicht alle $= 0$ sind, so daß

$$\sum_{j=1}^{n+1} \lambda_j v_j = \lambda_1 v_1 + \ldots + \lambda_{n+1} v_{n+1} = 0 \tag{10}$$

ist. Nach Voraussetzung gibt es ein Erzeugendensystem $(b_1, \ldots, b_n)$ von V und damit $\mu_{ij} \in K$, $i = 1, \ldots, n$, $j = 1, \ldots, n+1$, so daß

$$v_j = \mu_{1j} b_1 + \ldots + \mu_{nj} b_n = \sum_{i=1}^{n} \mu_{ij} b_i.$$

Damit ist

$$\sum_{j=1}^{n+1} \lambda_j v_j = \sum_{j=1}^{n+1} \lambda_j \left(\sum_{i=1}^{n} \mu_{ij} b_i \right) = \sum_{i=1}^{n} \left(\sum_{j=1}^{n+1} \mu_{ij} \lambda_j \right) b_i.$$

Es genügt also zu zeigen, daß es zu beliebig vorgegebenen μ_{ij} Werte λ_j gibt, die nicht alle $= 0$ sind, so daß

$$\sum_{j=1}^{n+1} \mu_{ij}\lambda_j = 0 \quad \text{für alle } i = 1, \ldots, n.$$

Das ist aber ein homogenes lineares Gleichungssystem mit n Gleichungen für $n+1$ Unbekannte, und das hat immer eine nicht-triviale Lösung, vgl. Lemma 1.

2 $\implies$ 1. Sei $(b_1, \ldots, b_r)$ ein linear unabhängiges r-tupel mit maximalem r. Das existiert nach Voraussetzung, und es ist $1 \leq r \leq n$. Sei

$$U = \text{Spann}(b_1, \ldots, b_r).$$

Ist $U = V$, so sind wir fertig. Andernfalls gibt es ein $b_{r+1} \in V \setminus U$. Nach Lemma 2 sind dann $(b_1, \ldots, b_{r+1})$ linear unabhängig im Widerspruch zur Maximalität von r. □

Satz 4 (Basissatz). *Sei V ein endlich erzeugter K-Vektorraum. Dann gilt:*

1. *V besitzt eine endliche Basis. Genauer: Ist M ein endliches Erzeugendensystem, so gibt es eine Basis $(b_1, \ldots, b_n)$ für V mit $b_i \in M$ für alle i.*
2. *Je zwei Basen von V haben dieselbe Anzahl von Elementen.*
3. *Sind $b_1, \ldots, b_r \in V$ linear unabhängig und* keine *Basis von V, so gibt es Vektoren $b_{r+1}, \ldots, b_n \in V$, so daß $b_1, \ldots, b_n$ eine Basis von V ist. (Ergänzungssatz)*

Bemerkung *Jeder* Vektorraum (auch wenn er nicht endlich erzeugt ist) besitzt eine Basis. Wir beweisen das nicht, weil wir es in der linearen Algebra fast ausschließlich mit dem endlich erzeugten Fall zu tun haben.

Beweis des Basissatzes. Ist $V = \{0\}$, so sind die Behauptungen trivial. Wir setzen deshalb $V \neq \{0\}$ voraus.

Zu 1. Sei $(b_1, \ldots, b_n)$ ein n-tupel linear unabhängiger Elemente von M mit maximalem n. Setze

$$U := \text{Spann}(b_1, \ldots, b_n).$$

Ist $M \subset U$, so folgt $V = \text{Spann}(M) \subset \text{Spann}(U) = U$, also $V = \text{Spann}(b_1, \ldots, b_r)$, und wir sind fertig. Andernfalls gibt es $b_{n+1} \in M \setminus U$. Dann sind $b_1, \ldots, b_{n+1}$ nach Lemma 2 linear unabhängig, im Widerspruch zur Maximalität von n. Dieser Fall kommt also nicht vor.

Zu 2. Sind B_1 und B_2 zwei Basen mit n_1 bzw. n_2 Elementen, und ist etwa $n_1 \leq n_2$, so ist B_1 ein Erzeugendensystem von V mit n_1 Elementen, und weil die n_2 Elemente von B_2 linear unabhängig sind, ist nach dem Schrankenlemma $n_2 \leq n_1$. Also folgt $n_1 = n_2$.

Zu 3. V sei erzeugt von n Elementen. Sind $(b_1, \ldots, b_r)$ linear unabhängig aber keine Basis, so sind sie also kein Erzeugendensystem. D.h.

$$U := \text{Spann}(b_1, \ldots, b_r) \neq V.$$

Dann gibt es $b_{r+1} \in V \setminus U$ und nach Lemma 2 sind $(b_1, \ldots, b_{r+1})$ linear unabhängig. Diesen Prozeß können wir fortsetzen, solange wir nicht bei einem Erzeugendensystem angekommen sind. Aber nach dem Schrankenlemma geht das allenfalls, solange $r < n$ ist. Also bricht der Erweiterungsprozeß ab, weil wir bei einem Erzeugendensystem angekommen sind. □

Beispiel 30. Sei V ein K-Vektorraum mit Basis $(b_1, \ldots, b_n)$ und $n \geq 2$. Dann ist auch

$$(b_1 + b_2, b_2, \ldots, b_n)$$

eine Basis, woraus Sie ersehen können, daß ein Vektorraum im allgemeinen nicht nur eine Basis besitzt!

□

Ist $(b_1, \ldots, b_n)$ ein *Erzeugendensystem* von V, so läßt sich jedes $v \in V$ schreiben als

$$v = \lambda_1 b_1 + \ldots + \lambda_n b_n$$

mit gewissen *Koeffizienten* $\lambda_i \in K$. Ist $(b_1, \ldots, b_n)$ eine *Basis* von V, so sind die *Koeffizienten* eindeutig bestimmt:

Satz 5 (Eindeutigkeitssatz). *Ist $(b_1, \ldots, b_n)$ eine Basis des K-Vektorraums V und sind $\lambda_1, \ldots, \lambda_n, \mu_1, \ldots, \mu_n \in K$ mit*

$$\lambda_1 b_1 + \ldots + \lambda_n b_n = \mu_1 b_1 + \ldots + \mu_n b_n,$$

so folgt

$$\lambda_1 = \mu_1, \ldots, \lambda_n = \mu_n.$$

Beweis. Aus der angenommenen Gleichheit folgt

$$(\lambda_1 - \mu_1) b_1 + \ldots + (\lambda_n - \mu_n) b_n = 0.$$

Weil die b_i linear unabhängig sind, folgt die Behauptung. □

Definition 15 (Dimension). Sei V ein K-Vektorraum. Wir definieren die Dimension von V

$$\dim V \in \{0, 1, 2, \ldots\} \cup \{\infty\}$$

wie folgt:

1. Ist $V = \{0\}$, so sei
$$\dim V = 0.$$
2. Ist $V \neq \{0\}$ endlich erzeugt mit Basis $(b_1, \ldots, b_n)$, so sei
$$\dim V = n.$$
3. Ist V nicht endlich erzeugt, so sei
$$\dim V = \infty.$$

Bemerkung: Ein Vektorraum V ist also genau dann endlich-dimensional, wenn er endlich erzeugt ist. Mit dem Begriff der Dimension ergibt sich aus dem Basissatz

Satz 6. *Ist V ein Vektorraum der Dimension $n < \infty$ und sind $b_1, \ldots, b_n \in V$ linear unabhängig, so bilden sie eine Basis von V.*

Beweis. Wären sie keine Basis, könnte man sie nach dem Basisergänzungssatz zu einer Basis ergänzen, die dann aber mehr als n Elemente enthielte. Widerspruch zu $\dim V = n$. □

Beispiel 31.

$$\dim K^n = n,$$
$$\dim \mathrm{Pol}(\mathbb{R}) = \infty.$$

Setzt man $\mathrm{Pol}_n(\mathbb{R}) = \{p(x) \in \mathrm{Pol}(\mathbb{R}) \mid \mathrm{Grad}\, p(x) \leq n\}$, so ist

$$\dim \mathrm{Pol}_n(\mathbb{R}) = n + 1.$$

Eine Basis ist gegeben durch die Monome $1, x, \ldots, x^n$.

$\square$

Satz 7 (Dimensionssatz). *Sei V ein Vektorraum und $n \in \mathbb{N}$. Dann gilt*

$$\dim V \leq n \iff \textit{Je } n+1 \textit{ Vektoren in } V \textit{ sind linear abhängig.} \tag{11}$$
$$\dim V = \infty \iff V \textit{ enthält eine unendliche linear unabhängige Teilmenge.} \tag{12}$$

Beweis. Zu (11) $\implies$. Schrankenlemma.

Zu (11) $\Leftarrow$. Schrankenlemma und Basissatz.

Zu (12) $\implies$. Nach Definition ist V nicht endlich erzeugt, d.h. es besitzt keine endliches Erzeugendensystem. Insbesondere ist $V \neq \{0\}$ und es gibt einen linear unabhängigen Vektor $b_1 \in V$. Sind $b_1, \ldots, b_r$ linear unabhängig, so ist nach Voraussetzung

$$U := \mathrm{Spann}(b_1, \ldots, b_r) \neq V$$

und durch Wahl eines $b_{r+1} \in V \backslash U$ verlängern wir die Kette der $b's$ linear unabhängig um ein weiteres Element. Auf diese Weise erhalten wir eine unendliche linear unabhängige Menge

$$M = \{b_i \mid i \in \mathbb{N}\}.$$

Zu (12) $\Leftarrow$. Enthält V eine unendliche linear unabhängige Menge, so enthält es auch endliche linear unabhängige Mengen von beliebiger Mächtigkeit. Nach dem Schrankenlemma ist V dann nicht endlich erzeugt. $\square$

Korollar 1 (Dimension von Unterräumen). *Sei $U \subset V$ ein Vektorunterraum des endlich-dimensionalen Vektorraums V. Dann ist auch U endlich-dimensional und*

$$\dim U \leq \dim V.$$

Gleichheit steht genau dann, wenn $U = V$ ist.

Bemerkung. Man bezeichnet

$$\dim V - \dim U$$

als die *Kodimension* des Unterraumes U und nennt Unterräume der Kodimension 1 auch *Hyperebenen.*

Beweis. Sei $n := \dim V$. Dann sind je $n+1$ Vektoren aus V linear abhängig. Also sind je $n+1$ Vektoren aus U linear abhängig. Nach dem Schrankenlemma wird U daher von n Elementen erzeugt. Eine Basis $(b_1, \ldots, b_m)$ von U ist entweder auch eine Basis von V, dann sind die Dimensionen gleich und $U = V$, oder sie läßt sich nach dem Basissatz zu einer Basis von V ergänzen, und dann ist $\dim U = m < n = \dim V$. $\square$

Korollar 2. *Seien $U, \tilde{U} \subset V$ zwei Vektorunterräume des endlich-dimensionalen Vektorraums V. Dann gilt*

$$\dim(U + \tilde{U}) = \dim U + \dim \tilde{U} - \dim(U \cap \tilde{U}).$$

Beweis. Sei $u_1, \ldots, u_k$ eine Basis von $U \cap \tilde{U}$. Wir ergänzen sie zu einer Basis

$$u_1, \ldots, u_k, u_{k+1}, \ldots, u_n \text{ von } U$$

und zu einer Basis

$$u_1, \ldots, u_k, \tilde{u}_{k+1}, \ldots, \tilde{u}_{\tilde{n}} \text{ von } \tilde{U}$$

Dann ist

$$(u_1, \ldots, u_k, u_{k+1}, \ldots, u_n, \tilde{u}_{k+1}, \ldots, \tilde{u}_{\tilde{n}})$$

ein Erzeugendensystem von $U + \tilde{U}$. Wir zeigen noch die lineare Unabhängigkeit. Dann ist

$$\dim(U + \tilde{U}) = n + \tilde{n} - k = \dim U + \dim \tilde{U} - \dim(U \cap \tilde{U}).$$

Aus

$$0 = \sum_{i=1}^{n} \lambda_i u_i + \sum_{i=k+1}^{\tilde{n}} \tilde{\lambda}_i \tilde{u}_i$$

folgt

$$\sum_{i=k+1}^{\tilde{n}} \tilde{\lambda}_i \tilde{u}_i = -\sum_{i=1}^{n} \lambda_i u_i.$$

Die linke Seite liegt in $\tilde{U}$, die rechte in U, also beide in $U \cap \tilde{U}$. Nach Konstruktion der Basis von $\tilde{U}$ ist dann aber die linke Seite $= 0$, also alle $\tilde{\lambda}_i = 0$ und ebenso alls $\lambda_i = 0$. □

Die wesentliche Leistung unserer theoretischen Überlegungen bisher sind die Begriffe *Vektorraum* und *Dimension.* Zur „Anwendung“ inbesondere unserer Einsichten über die Dimension kommen wir zurück auf frühere Beispiele.

Beispiel 32 (Homogene lineare Gleichungssysteme). Im Beispiel 13 haben wir festgestellt, daß die Lösungsmenge eines homogenen linearen Gleichungssystems ein Untervektorraum des K^n ist. Im ersten Abschnitt haben wir festgestellt, daß wir ein solches System durch den Gaußalgorithmus vereinfachen können. Danach bleiben (jedenfalls bei *homogenen* Systemen) r von Null verschiedene Zeilen in Zeilenstufenform übrig. Und damit zerfallen die Komponenten x_i der Lösungen $(x_1, \ldots, x_n)$ in zwei Klassen: die *Kopfvariablen*, die am Kopf einer dieser r Zeilen stehen, und die übrigen *freien* Variablen. Die r Kopfvariablen ergeben sich nach beliebiger Wahl der freien dann aus dem Gaußalgorithmus eindeutig.

Jetzt kommen wir zur Theorie: Wir können jeweils eine der „freien“ Komponenten $= 1$ und alle anderen $= 0$ wählen und dann die Kopfvariablen bestimmen. Das gibt $n - r$ verschiedene spezielle Lösungen des homogenen Systems, die sogar linear unabhängig sind. Das Argument für die lineare Unabhängigkeit ist dasselbe wie für die lineare Unabhängigkeit der Standardbasis $(e_1, \ldots, e_n)$.

Weil der Lösungsraum ein Untervektorraum ist, ist jede Linearkombination dieser $n - r$ Lösungen wieder eine Lösung. Und mit einer geeigneten Linearkombination können wir jede Wahl der „freien“ Variablen realisieren, d.h. jede Lösung ist eine Linearkombination der obigen speziellen $n - r$ Lösungen.

Wir schreiben das auf für den Spezialfall, daß in der Zeilenstufenform

$$a_{ii} \neq 0 \text{ für } 1 \leq i \leq r,$$

d.h. daß $x_1, \ldots, x_r$ die Kopfvariablen sind. Dann kann man also $x_{r+1}, \ldots, x_n$ frei wählen und wir bezeichnen mit $x^{(r+i)}, i \in \{1, \ldots, n-r\}$ die Lösung des homogenen Systems mit

$$x_{r+j}^{(r+i)} = \begin{cases} 1 & \text{für } j = i, \\ 0 & \text{sonst.} \end{cases}$$

Die speziellen Lösungen sehen also so aus

$$x^{(r+1)} = \begin{pmatrix} * \\ \vdots \\ * \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, x^{(r+2)} = \begin{pmatrix} * \\ \vdots \\ * \\ 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \ldots, x^{(n)} = \begin{pmatrix} * \\ \vdots \\ * \\ 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}.$$

Sie sehen unmittelbar die lineare Unabhängigkeit. Linearkombinationen dieser Lösungen sind wieder Lösungen, und mit

$$x = \lambda_{r+1} x^{(r+1)} + \ldots + \lambda_n x^{(n)} = \begin{pmatrix} * \\ \vdots \\ * \\ \lambda_{r+1} \\ \vdots \\ \lambda_n \end{pmatrix}$$

erhält man *die eindeutig bestimmte* Lösung mit $x_{r+1} = \lambda_{r+1}, \ldots, x_n = \lambda_n$. Die $x^{(r+i)}$ sind also eine Basis des Lösungsraumes, und dieser hat die Dimension $n - r$.

□

Als Konsequenz aus der Definition der Dimension (insbesondere aus der Tatsache, daß je zwei Basen dieselbe Mächtigkeit haben) ergibt sich damit der

Satz 8 (Rang linearer Gleichungssysteme). *Wendet man auf ein lineares Gleichungssystem den Gaußalgorithmus an, so ist die Zahl r der verbleibenden „relevanten“ Zeilen unabhängig von der Durchführung des Algorithmus. Es ist*

$$n - r = \textit{Dimension des Lösungsraumes des zugehörigen homogenen Systems.}$$

Man nennt r den Rang *des linearen Gleichungssystems.*

Ein weiteres Ergebnis ist, daß man *alle* Lösungen eines homogenen linearen Gleichungssystems vom Rang r durch Linearkombination aus beliebigen $n - r$ linear unabhängigen Lösungen erhält, *egal, wie man diese letzteren gefunden hat.*

Beispiel 33 (Lineare Differentialgleichungen). Im Beispiel 14 haben wir gesehen, daß die Menge der Lösungen der Differentialgleichung

$$y'' + y = 0 \tag{13}$$

ein Vektorunterraum des Funktionenraumes $\mathbb{R}^{\mathbb{R}}$ aller Abbildungen von $\mathbb{R}$ nach $\mathbb{R}$ ist. In der Theorie der Differentialgleichungen zeigt man, daß es zu einem gegebenen Paar $(\lambda_1, \lambda_2) \in \mathbb{R}^2$

genau eine Lösung $y(x)$ der Differentialgleichung mit

$$y(0) = \lambda_1 \text{ und } y'(0) = \lambda_2$$

gibt. Bezeichnet man mit $y_1(x)$ bzw. mit $y_2(x)$ die Lösungen mit

$$y(0) = 1, y'(0) = 0$$

bzw. mit

$$y(0) = 0, y'(0) = 1,$$

so ist

$$y(x) = \lambda_1 y_1(x) + \lambda_2 y_2(x)$$

gerade die Lösung mit den Anfangswerten $y(0) = \lambda_1, y'(0) = \lambda_2$. Aber

$$y_1(x) = \cos x, \quad y_2(x) = \sin x,$$

und diese beiden Funktionen sind linear unabhängig. Der Lösungsraum von (13) ist also ein zweidimensionaler Untervektorraum des (unendlich-dimensionalen) Vektorraums $\mathbb{R}^{\mathbb{R}}$.

Dies ist ein Spezialfall folgender allgemeinen Situation: Die homogene lineare Differentialgleichung n-ter Ordnung

$$y^{(n)} + a_1(x)y^{(n-1)} + \ldots + a_n(x)y = 0 \tag{14}$$

mit stetigen Funktionen $a_i(x)$ auf einem Intervall I hat genau eine Lösung $y(x)$ mit vorgegebenen Anfangswerten

$$y(x^*) = \lambda_1, \ldots, y^{(n-1)}(x^*) = \lambda_n.$$

Dabei ist $x^* \in I$ ein fester Punkt. Dies ist ein schwierigeres Resultat aus der Theorie der Differentialgleichungen. Die Menge der Lösungen bildet einen Vektorunterraum des Raums aller reellen Funktionen auf I. Das ist leicht nachzuweisen. Wählt man nun spezielle Lösungen $y_i, 1 \leq i \leq n$ mit

$$y_i^{(j-1)}(x) = \begin{cases} 1 & \text{für } i = j, \\ 0 & \text{sonst,} \end{cases}$$

so sind das n *linear unabhängige* Lösungen (Beweis?), aus denen man wieder jede (Anfangsbedingung, also jede) Lösung linear kombinieren kann. Der Lösungsraum von (14) ist also ein n-dimensionaler Vektorraum. Man erhält alle Lösungen durch Linearkombination, wenn man (irgendwie) n linear unabhängige gefunden hat.

Dieses Ergebnis unserer Theorie erhält für die *Differential*gleichungen deshalb eine besondere Bedeutung, weil es keinen Algorithmus gibt, der einem Lösungen produziert. Man muß vielmehr für jede solche homogene lineare Differentialgleichung, z.B. für

$$y'' - \frac{2x}{1-x^2}\, y' + \frac{20}{1-x^2}\, y = 0,$$

mit irgendwelchen Tricks mühevoll nach Lösungen suchen. Natürlich ist es dann von fundamentaler Bedeutung zu wissen, wann man genug gefunden hat!

□

4 Quotientenvektorräume

Wir schieben hier einen kurzen Abschnitt über die Konstruktion von Quotientenvektorräumen ein. Seien V ein K-Vektorraum und $U \subset V$ ein Untervektorraum. Dann definieren wir eine Äquivalenzrelation auf V durch

$$v \sim w :\Longleftrightarrow v - w \in U.$$

Das *ist* eine Äquivalenzrelation, denn es gelten die dafür nötigen drei Axiome:

$$\begin{gathered} v \sim v, \\ v \sim w \implies w \sim v, \\ v \sim w \text{ und } w \sim z \implies v \sim z. \end{gathered}$$

Wir bezeichnen die Äquivalenzklasse von v mit $[v]$. Es ist also

$$[v] = v + U = \{v + u \,|\, u \in U\}$$

der affine Unterraum durch v mit Richtungsvektorraum U.

Definition 16. Seien V ein K-Vektorraum und $U \subset V$ ein Untervektorraum. Dann definieren wir den *Quotientenvektorraum* V/U als die Menge der Restklassen

$$V/U := \{[v] \,|\, v \in V\}$$

mit den durch

$$\begin{aligned} [v] + [w] &:= [v + w] \\ \lambda[v] &:= [\lambda v] \end{aligned}$$

gegebenen Verknüpfungen.

Beweis. 1. müssen wir zeigen, daß die Rechenoperationen wohldefiniert sind.

Wir müssen zeigen, daß für $v_1, v_2, w_1, w_2 \in V$ gilt

$$[v_1] = [v_2] \text{ und } [w_1] = [w_2] \implies [v_1 + w_1] = [v_2 + w_2]$$

und Entsprechendes für die Skalarmultiplikation. Aber aus

$$v_1 - v_2 \in U \text{ und } w_1 - w_2 \in U$$

folgt offensichtlich

$$(v_1 + w_1) - (v_2 + w_2) = (v_1 - v_2) + (w_1 - w_2) \in U.$$

Ebenso schließt man für die Skalarmultiplikation.

2. sind noch die Vektorraumaxiome nachzuprüfen. Aber das unterschlagen wir hier, weil es so langweilig ist. Wir stellen nur fest, daß

$$0_{V/U} = [0] = U,$$

also

$$[v] = 0 \iff v \in U.$$

□

Beispiel 34. Sei $V := \{f : [0,1] \to \mathbb{R} \,|\, f \text{ stetig}\}$ der Vektorraum der stetigen Funktionen auf dem Intervall $[0,1]$. Zum Beispiel ist die Funktion $\mathbf{1} : [0,1] \to \mathbb{R}, x \mapsto 1$ in V. Die Menge

$$U := \{f \in V \,|\, \int_0^1 f(x)dx = 0\}$$

ist ein Untervektorraum von V (Beweis?), und wir behaupten

$$([\mathbf{1}]) \text{ ist eine Basis von } V/U.$$

Weil $\int_0^1 1dx = 1 \neq 0$, ist $\mathbf{1} \notin U$, also $[\mathbf{1}] \neq 0$. Also ist $([\mathbf{1}])$ linear unabhängig. Bleibt zu zeigen, das $([\mathbf{1}])$ ein Erzeugendensystem ist, d.h. daß es zu jedem $f \in V$ ein $\lambda \in \mathbb{R}$ gibt, für das

$$[f] = \lambda[\mathbf{1}].$$

Das ist aber äquivalent zu

$$f - \lambda\mathbf{1} \in U$$

oder

$$0 = \int_0^1 (f(x) - \lambda)dx = \int_0^1 f(x)dx - \lambda \int_0^1 dx = \int_0^1 f(x)dx - \lambda.$$

Also können (und müssen) wir $\lambda = \int_0^1 f(x)dx$ wählen.

□

Satz 9 (Quotientenvektorraum). *Seien V ein endlich-dimensionaler K-Vektorraum und $U \subset V$ ein Untervektorraum. Dann gilt*

$$\dim V/U = \dim V - \dim U.$$

Ist $1 \leq r \leq n$ und $(v_1, \ldots, v_n)$ eine Basis von V, so daß $v_{r+1}, \ldots, v_n$ eine Basis von U ist, so ist

$$([v_1], \ldots, [v_r])$$

eine Basis von V/U.

Beweis. Wir zeigen zunächst die zweite Behauptung. Dazu genügt zu zeigen, daß $([v_1], \ldots, [v_r])$ linear unabhängig sind und V/U aufspannen. Nun bedeutet aber

$$\lambda_1[v_1] + \ldots + \lambda_r[v_r] = 0_{V/U},$$

einfach

$$\lambda_1 v_1 + \ldots + \lambda_r v_r \in U,$$

d.h. es gibt $\lambda_{r+1}, \ldots \lambda_n$, so daß

$$\lambda_1 v_1 + \ldots + \lambda_r v_r = \lambda_{r+1} v_{r+1} + \ldots + \lambda_n v_n.$$

Aus der Basiseigenschaft von $(v_1, \ldots, v_n)$ folgt, daß alle $\lambda_i = 0$.

Ist weiter $v \in V$, etwa

$$v = \sum_{i=1}^n \lambda_i v_i,$$

so ist

$$[v] = \sum_{i=1}^n \lambda_i [v_i] = \sum_{i=1}^r \lambda_i [v_i],$$

weil $v_{r+i} \in U$, also $[v_{r+i}] = 0$. Also ist $([v_1], \ldots, [v_r])$ ein Erzeugendensystem von V/U.

Die erste Behauptung folgt aus der zweiten: Man wähle eine Basis $(v_{r+1}, \ldots, v_n)$ für U und ergänze sie zu einer Basis $(v_1, \ldots, v_n)$ von V (Basisergänzungssatz).

□

5 Lineare Abbildungen

Wir betrachten nun Abbildungen zwischen Vektorräumen, welche die Struktur der Vektorräume erhalten:

Definition 17. Seien V, W zwei K-Vektorräume. Eine Abbildung

$$f : V \to W$$

heißt *linear*, wenn für alle $v, v_1, v_2 \in V$ und $\lambda \in K$ folgende Bedingungen gelten:

$$\begin{aligned} f(v_1 + v_2) &= f(v_1) + f(v_2) && \text{(Additivität)}, \\ f(\lambda v) &= \lambda f(v) && \text{(Homogenität)}. \end{aligned}$$

Beispiel 35. Wir betrachten $V = W = K^1 = K$ und $a, b \in K$. Die Abbildung

$$g : K \to K, v \mapsto av + b$$

ist genau dann linear, wenn $b = 0$.

$\square$

Beispiel 36. Die Identität

$$\mathrm{Id} : V \to V, v \mapsto v$$

und die Nullabbildung

$$N : V \to W, v \mapsto 0$$

sind lineare Abbildungen.

$\square$

Beispiel 37. Die Abbildung

$$f : \mathbb{R}^3 \to \mathbb{R}^2$$

mit

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

ist linear, denn es ist

$$\begin{aligned} f(\begin{pmatrix} x_1 + x_2 \\ y_1 + y_2 \\ z_1 + z_2 \end{pmatrix}) &= \begin{pmatrix} 2(x_1 + x_2) - (y_1 + y_2) - 3(z_1 + z_2) \\ (x_1 + x_2) + (y_1 + y_2) \end{pmatrix} \\ &= \begin{pmatrix} 2x_1 - y_1 - 3z_1 \\ x_1 + y_1 \end{pmatrix} + \begin{pmatrix} 2x_2 - y_2 - 3z_2 \\ x_2 + y_2 \end{pmatrix} \\ &= f(\begin{pmatrix} x_1 \\ y_1 \\ z_1 \end{pmatrix}) + f(\begin{pmatrix} x_2 \\ y_2 \\ z_2 \end{pmatrix}) \end{aligned}$$

und entsprechend

$$f(\lambda \begin{pmatrix} x \\ y \\ z \end{pmatrix}) = \lambda f(\begin{pmatrix} x \\ y \\ z \end{pmatrix}).$$

$\square$

Beispiel 38. Seien $J \subset \mathbb{R}$ ein Interval und $k \in \mathbb{N}$ eine positive natürliche Zahl. Die Mengen

$$\begin{aligned} C^0(J) &:= \{f : J \to \mathbb{R} \,|\, f \text{ ist stetig}\}, \\ C^k(J) &:= \{f : J \to \mathbb{R} \,|\, f \text{ ist } k\text{-mal differenzierbar und } f^{(k)} \in C^0(J)\} \end{aligned}$$

sind Untervektorräume von $\mathbb{R}^J$. Die Abbildungen

$$D : C^{(k)}(J) \to C^{(k-1)}(J), f \mapsto f'$$

und, für $a, b \in J$,

$$I : C^0(J) \to \mathbb{R}, f \mapsto \int_a^b f(x)dx$$

sind lineare Abbildungen. Das besagt gerade, daß man Summen gliedweise differenzieren bzw. integrieren kann, und daß konstante Faktoren erhalten bleiben.

$\square$

Beispiel 39. Seien V ein K-Vektorraum und $U \subset V$ ein Vektorunterraum. Dann ist die sogenannte *kanonische Projektion*

$$\pi : V \to V/U, v \mapsto \pi(v) = [v]$$

eine lineare Abbildung.

Beweis. Es ist

$$\pi(v + w) = [v + w] = [v] + [w] = \pi(v) + \pi(w)$$

nach Definition der Addition in V/U. Entsprechendes gilt für die Skalarmultiplikation. $\square$

$\square$

Beispiel 40 (Komposition linearer Abbildungen). Seien U, V, W drei K-Vektorräume und

$$g : U \to V, \quad f : V \to W$$

lineare Abbildungen. Dann ist auch die Komposition

$$f \circ g : U \to W, u \mapsto f \circ g(u) := f(g(u))$$

eine lineare Abbildung.

$\square$

Lemma 4. *Sei $f : V \to W$ linear. Dann gilt:*

$$f(0) = 0$$

Weiter sind die beiden folgenden Aussagen äquivalent:

1. *f ist injektiv*
2. *Für alle $v \in V$ gilt*
$$f(v) = 0 \iff v = 0.$$

Beweis. Zunächst ist

$$f(0_V) = f(0_K 0_V) = 0_K f(0_V) = 0_W.$$

Daraus folgt die erste Behauptung. Ist weiter f injektiv, so folgt aus $f(v) = 0 = f(0)$, daß $v = 0$. Ist umgekehrt $f(v) = 0$ nur für v=0, so folgt aus

$$f(v_1) = f(v_2),$$

daß $0 = f(v_1) - f(v_2) = f(v_1 - v_2)$, also $v_1 - v_2 = 0$. $\square$

Definition 18 (Endomorphismen, Isomorphismen). Seien V, W zwei K-Vektorräume.

1. Eine lineare Abbildung $f : V \to V$ von V in sich nennt man auch einen *Endomorphismus.*
2. Eine *bijektive* lineare Abbildung $f : V \to W$ heißt ein *Isomorphismus* (von Vektorräumen).
3. Einen Isomorphismus $f : V \to V$ nennt man auch einen *Automorphismus.*
4. V heißt *isomorph zu* W, falls es einen Isomorphismus $f : V \to W$ gibt. Notation:
$$V \cong W.$$

Bemerkung. Ist $f : V \to W$ ein Isomorphismus, so ist auch die Umkehrabbildung
$$f^{-1} : W \to V$$
eine *lineare* Abbildung. Daher ist Isomorphie eine Äquivalenzrelation.

Satz 10 (Existenz und Einzigkeit linearer Abbildungen). *Seien V, W zwei K-Vektorräume, $(v_1, \ldots, v_n)$ eine Basis von V und $w_1, \ldots, w_n \in W$ beliebige Vektoren. Dann gibt es genau eine lineare Abbildung $f : V \to W$ mit*
$$f(v_i) = w_i \quad \textit{für alle } i \in \{1, \ldots, n\} \tag{15}$$

Beweis. Beweis der Einzigkeit. Jedes $v \in V$ besitzt eine eindeutige Darstellung
$$v = \lambda_1 v_1 + \ldots + \lambda_n v_n.$$
Wenn es eine lineare Abbildung f mit der angegebenen Eigenschaft überhaupt gibt, so muß also
$$f(v) = f(\lambda_1 v_1 + \ldots + \lambda_n v_n) = \lambda_1 f(v_{1)} + \ldots + \lambda_n f(v_n) = \lambda_1 w_1 + \ldots + \lambda_n w_n$$
sein. Daher gibt es *höchstens eine* solche lineare Abbildung f.

Beweis der Existenz. Dazu definieren wir für $v = \lambda_1 v_1 + \ldots + \lambda_n v_n$
$$f(v) := \lambda_1 w_1 + \ldots + \lambda_n w_n \tag{16}$$
Das ist die einzige Chance, die wir haben, ein solches f zu finden! Wir müssen zeigen, daß (16) tatsächlich eine lineare Abbildung liefert. Die Eigenschaft (15) ist dann klar. Nun ist aber
$$\lambda v = \lambda\lambda_1 v_1 + \ldots + \lambda\lambda_n v_n$$
und daher
$$f(\lambda v) = \lambda\lambda_1 w_1 + \ldots + \lambda\lambda_n w_n = \lambda(\lambda_1 w_1 + \ldots + \lambda_n w_n) = \lambda f(v)$$
Das beweist die Homogenität von f. Die Additivität zeigt man ähnlich. □

Satz 11 (Fundamentalsatz über endlich-dimensionale Vektorräume). *Seien V und W zwei endlich-dimensionale K-Vektorräume. Dann gilt*
$$\dim V = \dim W \iff V \cong W.$$

Beweis. Zu $\implies$. Sei $n := \dim V = \dim W$.

Ist $n = 0$, so ist $V = \{0\} = W$ und es ist nichts zu zeigen.

Sei also $n > 0$. Seien $(v_1, \ldots, v_n)$ und $(w_1, \ldots, w_n)$ Basen von V bzw. W. Dann gibt es genau eine lineare Abbildung $f : V \to W$ mit

$$f(v_i) = w_i \text{ für alle } i$$

und genau eine lineare Abbildung $g : W \to V$ mit

$$g(w_i) = v_i \text{ für alle } i.$$

Dann ist $g \circ f : V \to V$ eine lineare Abbildung mit

$$g \circ f(v_i) = v_i \text{ für alle } i.$$

Aber $\mathrm{Id} : V \to V$ ist ebenfalls eine lineare Abbildung mit dieser Eigenschaft. Aus der Eindeutigkeitsaussage von Satz 10 folgt daher

$$g \circ f = \mathrm{Id} : V \to V.$$

Genauso folgt $f \circ g = \mathrm{Id} : W \to W$. Daher ist f bijektiv, also ein Isomorphismus.

Zu $\Leftarrow$. Seien $f : V \to W$ ein Isomorphismus, $(v_1, \ldots, v_n)$ eine Basis von V und

$$w_i := f(v_i), \qquad i \in \{1, \ldots, n\}.$$

Es genügt zu zeigen, daß $(w_1, \ldots, w_n)$ eine Basis von W ist. Zunächst folgt aus

$$0 = \lambda_1 w_1 + \ldots + \lambda_n w_n = \lambda_1 f(v_1) + \ldots + \lambda_n f(v_n) = f(\lambda_1 v_1 + \ldots + \lambda_n v_n)$$

durch Anwendung der linearen Abbildung f^{-1}:

$$0 = f^{-1}(0) = \lambda_1 v_1 + \ldots + \lambda_n v_n.$$

Aus der Basiseigenschaft der v_i folgt $\lambda_i = 0$ für alle i, und die w_i sind linear unabhängig.

Sei weiter $w \in W$. Dann gibt es $v \in V$ mit $f(v) = w$. Sei

$$v = \lambda_1 v_1 + \ldots + \lambda_n v_n.$$

Dann ist

$$w = f(\lambda_1 v_1 + \ldots + \lambda_n v_n) = \lambda_1 w_1 + \ldots + \lambda_n w_n.$$

Daher ist $W = \mathrm{Spann}(w_1, \ldots, w_n)$ und die w_i bilden eine Basis von W. □

Bemerkung. Da ein endlich-dimensionaler Vektorraum im allgemeinen viele Basen hat, von denen keine besonders ausgezeichnet ist, gibt es zwischen zwei (endlich) gleich-dimensionalen Vektorräumen V, W im allgemeinen viele Isomorphismen, von denen aber keiner besonders ausgezeichnet ist. Man sagt, V und W sind isomorph, aber nicht *kanonisch isomorph*.

Andrerseits hat der Vektorraum K^n eine ausgezeichnete (kanonische) Basis $(e_1, \ldots, e_n)$, vgl. Beispiel 16. Ist dann $(v_1, \ldots, v_n)$ eine Basis des (abstrakten) Vektorraumes V, so gibt es einen eindeutig bestimmten Isomorphismus $f : K^n \to V$, unter dem sich diese Basen entsprechen, nämlich

$$f(\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}) = \sum_{i=1}^{n} x_i v_i.$$

Definition 19 (Koordinatensystem). Sei $\mathbf{v} = (v_1, \ldots, v_n)$ eine Basis des K-Vektorraums V. Dann heißt der Isomorphismus

$$\Phi_{\mathbf{v}}(\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}) = \sum_{i=1}^{n} x_i v_i,$$

das zur Basis $\mathbf{v}$ gehörige *Koordinatensystem* und man nennt $\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ den *Koordinatenvektor* von

$$v = \sum_{i=1}^{n} x_i v_i$$

bezüglich der Basis $\mathbf{v}$. Der Koordinatenvektor von v ist also $\Phi_{\mathbf{v}}^{-1}(v)$, und es gilt

$$\Phi_{\mathbf{v}}(e_j) = v_j. \tag{17}$$

Definition 20 (Kern und Bild). Sei $f : V \to W$ eine lineare Abbildung von K-Vektorräumen.

$$\operatorname{Kern} f := \{v \in V \mid f(v) = 0\}$$

heißt *der Kern* von f.

$$\operatorname{Bild} f := f(V) := \{f(v) \in W \mid v \in V\}$$

heißt *das Bild* von f.

Lemma 5. *Sei $f : V \to W$ eine lineare Abbildung von K-Vektorräumen. Dann sind* $\operatorname{Kern} f$ *und* $\operatorname{Bild} f$ *Untervektorräume von V bzw. W. Ist V oder W endlich-dimensional, so ist auch* $\operatorname{Bild} f$ *endlich-dimensional.*

Beweis. Die Unterraum-Eigenschaft ist leicht zu zeigen. Ist $\dim W < \infty$, so auch der Unterraum $\operatorname{Bild} f \subset W$. Ist schließlich $\dim V < \infty$ und $v_1, \ldots, v_n$ eine Basis von V, so ist

$$\operatorname{Bild} f = \operatorname{Spann}(f(v_1), \ldots, f(v_n)),$$

also endlich erzeugt und deshalb endlich-dimensional. □

Satz 12 (Homomorphiesatz für Vektorräume). *Sei $f : V \to W$ linear. Dann definiert*

$$\underline{f}([v]) := f(v), \quad v \in V$$

eine lineare Abbildung $\underline{f} : V/\operatorname{Kern} f \to \operatorname{Bild} f$. Diese ist ein Isomorphismus.

Beweis. Zunächst zeigen wir, daß $\underline{f}$ wohl-definiert ist.

Ist nämlich $[v_1] = [v_2]$, so ist $v_1 - v_2 \in \operatorname{Kern} f$ und deshalb

$$\underline{f}([v_2]) = f(v_2) = f(v_1 + (v_2 - v_1)) = f(v_1) = \underline{f}([v_1]).$$

Durch Nachrechnen zeigt man, daß $\underline{f}$ linear ist. Die Surjektivität folgt aus der Definition von $\operatorname{Bild} f$, und schließlich ist

$$\underline{f}([v_1]) = \underline{f}([v_2]) \iff f(v_1) = f(v_2) \iff v_1 - v_2 \in \operatorname{Kern} f \iff [v_1] = [v_2].$$

□

Definition 21 (Rang einer linearen Abbildung). Ist $f : V \to W$ linear und $\operatorname{Bild} f$ endlich-dimensional, so heißt

$$\operatorname{Rang} f := \dim \operatorname{Bild} f$$

der Rang der linearen Abbildung f.

Satz 13 (Dimensionsformel für lineare Abbildungen). *Sei $f : V \to W$ eine lineare Abbildung von endlich-dimensionalen K-Vektorräumen. Dann gilt:*

$$\dim V = \dim \operatorname{Kern} f + \dim \operatorname{Bild} f = \dim \operatorname{Kern} f + \operatorname{Rang} f. \tag{18}$$

Beweis. Die Formel folgt aus Satz 12 in Verbindung mit Satz 9:

$$\dim \operatorname{Bild} f = \dim V / \operatorname{Kern} f = \dim V - \dim \operatorname{Kern} f.$$

□

Korollar 3. *Seien V, W zwei K-Vektorräume gleicher endlicher Dimension. Sei $f : V \to W$ linear. Dann sind folgende Aussagen äquivalent:*

1. *f ist ein Isomorphismus,*
2. *f ist injektiv,*
3. $\operatorname{Kern} f = \{0\}$,
4. *f ist surjektiv.*

Beweis. 1. $\implies$ 2. Trivial.

2. $\implies$ 3. Trivial.

3. $\implies$ 4. Ist $\operatorname{Kern} f = \{0\}$, so ist $\dim \operatorname{Kern} f = 0$, also nach der Dimensionsformel $\dim \operatorname{Bild} f = \dim V = \dim W$ und daher nach Korollar 1 $\operatorname{Bild} f = W$.

4. $\implies$ 1. Ist f surjektiv, so ist $\operatorname{Rang} f = \dim W = \dim V$, also $\dim \operatorname{Kern} f = 0$, d.h. $\operatorname{Kern} f = \{0\}$. Aus Lemma 4 folgt, daß f auch injektiv, also ein Isomorphismus ist. □

6 Matrizen als lineare Abbildungen, lineare Gleichungssysteme

In diesem Abschnitt konkretisieren wir den Begriff linearer Abbildungen, um die Theorie linearer Gleichungssysteme oder linearer Differentialgleichungssysteme einfacher zu formulieren.

Definition 22 (Matrix). Seien $m, n \in \mathbb{N}$ positive natürliche Zahlen, K ein Körper und sei $a_{ij} \in K$ für alle $i \in \{1, \ldots, m\}$ und $j \in \{1, \ldots, n\}$. Dann heißt das „rechteckige Schema"

$$A = \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{pmatrix}$$

eine (m, n)-*Matrix über* K. Man schreibt dafür auch

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

oder kurz $A = (a_{ij})$, wenn die Laufbereiche der Indizes klar sind.

Bemerkung. Die Definition ersetzt den Begriff „Matrix“ durch den mathematisch nicht sehr exakten Begriff „rechteckiges Schema“. Formal exakt ist eine (m, n)-Matrix eine Abbildung

$$A : \{1, \ldots, m\} \times \{1, \ldots, n\} \to K, (i, j) \mapsto a_{ij}.$$

Desungeachtet ist die Notation der Werte dieser Abbildung in der angegebenen Form sehr übersichtlich und nützlich.

Beispiel 41 (Matrizen als lineare Abbildungen). Eine Matrix A liefert eine lineare Abbildung

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

wie folgt: Für $x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in K^n$ sei

$$f_A(x) := \begin{pmatrix} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n \end{pmatrix}. \tag{19}$$

Statt $f_A(x)$ schreibt man oft auch einfach Ax und nennt das aus später näher zu erläuternden Gründen *das Produkt* von A mit x.

Beachten Sie:

$$\boxed{\text{Die Spalten von } A \text{ sind gerade die Bilder } f_A(e_j) \text{ der Basisvektoren } e_j \in K^n.}$$

Ist $f : K^n \to K^m$ irgendeine lineare Abbildung und schreibt man die Bilder $f(e_j)$ in die Spalten einer Matrix A, so bildet f_A die e_j also auf dieselben Vektoren ab wie f. Nach Satz 10 folgt also $f = f_A$. Das bedeutet, daß nicht nur jede (m, n)-Matrix eine lineare Abbildung $f_A : K^n \to K^m$ induziert, sondern daß man auf diese Weise auch wirklich *jede* lineare Abbildung $K^n \to K^m$ erhält:

$\square$

Satz 14. *Die linearen Abbildungen von K^n in K^m sind genau die f_A mit beliebiger (m, n)-Matrix A.*

Beispiel 42 (Matrizen und lineare Gleichungssysteme). Ein lineares Gleichungssystem (3) definiert eine Matrix $A = (a_{ij})$, die man auch seine Systemmatrix nennt. Damit können wir das System *in Matrixschreibweise* so schreiben:

$$\begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix}.$$

Mit $b := \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix}$ wird das kurz

$$Ax = b.$$

Beim Gaußalgorithmus interessiert uns eigentlich nur die Systemmatrix, die man um eine Spalte, nämlich die rechte Seite, den sogenannten Zielvektor, erweitert hat. Also kann man sich die x_i sparen und den Gaußalgorithmus auf die erweiterte Systemmatrix anwenden. Wichtiger als die Schreibweise und schreibtechnische Vereinfachungen ist aber die inhaltliche Interpretation. Das lineare Gleichungssystem (3) definiert eine lineare Abbildung

$$f_A : K^n \to K^m,$$

und einen Vektor $b \in K^m$. Gesucht sind alle $x \in K^n$, die durch f_A auf b abgebildet werden.

Der mit dem Gaußalgorithmus ermittelte Rang eines linearen Gleichungssystems war $n-$ die Dimension des Lösungsraumes der homogenen Gleichung, also gleich dem Rang von f_A. Man nennt das auch den Rang der Matrix A, vgl. auch Satz 19. Den Rang einer Matrix kann man also ermitteln, indem man sie mit dem Gaußalgorithmus auf Zeilenstufenform bringt. Die Zahl der wesentlichen Zeilen ist der Rang der Matrix. Merken Sie sich:

$$\boxed{\text{Rang}\, A := \text{Rang}\, f_A = \text{Zahl der wesentlichen Zeilen „nach" Gaußalgorithmus.}}$$

□

Eine bis auf die Rangüberlegungen ähnliche Situation hat man bei linearen *Differentialgleichungen*:

Beispiel 43. Sei $I \subset \mathbb{R}$ ein Intervall, seien $b, a_1, \ldots, a_n : I \to \mathbb{R}$ stetig und bezeichne

$$C^n(I) := \{f : I \to \mathbb{R} \,|\, f \text{ ist } n\text{-mal stetig differenzierbar}\}$$

Dann ist

$$L : C^n(I) \to C^0(I), y \mapsto y^{(n)} + a_1 y^{(n-1)} + \ldots + a_n y$$

eine lineare Abbildung und

$$L(y) = b$$

ist eine *lineare Differentialgleichung n-ter Ordnung*. Ist $b = 0$ so nennt man diese *homogen*.

□

Wir nehmen das zum Anlaß für einen kurzen Einschub:

7 Intermezzo: Lösungsstruktur linearer Gleichungen.

Lineare Gleichungssysteme wie lineare Differentialgleichungen sind spezielle Fälle von „abstrakten" linearen Gleichungen der Form:

$$f(x) = w,$$

wobei $f : V \to W$ eine lineare Abbildung zwischen Vektorräumen ist. Die Gleichung

$$f(x) = 0,$$

nennt man *die zugehörige homogene Gleichung.* Aus der Linearität allein ergeben sich Strukturaussagen über die Lösungsmenge.

Homogene lineare Gleichungen $f(x) = 0$**.** Homogene lineare Gleichungen haben immer eine Lösung, nämlich die sogenannte triviale Lösung $x = 0$, weil lineare Abbildungen 0 auf 0 abbilden. Genauer ist die Lösungsmenge

$$f^{-1}(\{0\}) = \operatorname{Kern} f$$

ein Vektorunterraum von V. Linearkombinationen von Lösungen sind also wieder Lösungen. Ist V endlich-dimensional, so auch der Lösungsraum und

$$\dim \operatorname{Kern} f = \dim V - \operatorname{Rang} f.$$

Auch im Falle $\dim V = \infty$ kann der Lösungsraum endlich-dimensional sein, wie das Beispiel homogenener linearer Differentialgleichungen zeigt.

Inhomogene lineare Gleichungen $f(x) = w$**.** In diesem Fall kann die Lösungsmenge $f^{-1}(\{w\})$ leer sein, nämlich dann, wenn $\operatorname{Bild} f \neq W$ und w eben nicht im $\operatorname{Bild} f$ liegt. Ist allerdings

$$\operatorname{Rang} f = \dim W < \infty,$$

so folgt $\operatorname{Bild} f = W$, und die Gleichung ist für *jedes* $w \in W$ lösbar.

Falls $f(x) = w$ lösbar ist und z.B. x_S eine Lösung ist, man nennt das oft eine *spezielle* oder *partikuläre* Lösung, so ist der gesamte Lösungsraum gegeben durch

$$f^{-1}(\{w\}) = x_S + \operatorname{Kern} f = \{x_S + x \mid x \in \operatorname{Kern} f\}. \tag{20}$$

Man sagt

Der Lösungsraum der inhomogenen Gleichung ist entweder leer oder die Summe aus einer speziellen Lösung und dem Lösungsraum der homogenen Gleichung.

Beweis. Für $x \in \operatorname{Kern} f$ ist offenbar

$$f(x_S + x) = f(x_S) + f(x) = w + 0 = w.$$

Also ist $x_S + x$ eine Lösung. Aber in dieser Form erhält man *jede* Lösung. Ist nämlich $f(x_1) = w$, so ist $x_1 = x_S + (x_1 - x_S)$ und

$$f(x_1 - x_S) = f(x_1) - f(x_S) = w - w = 0.$$

□

Definition 23. Sei V ein Vektorraum und $A \subset V$. Dann heißt A ein *affiner Unterrraum von* V, wenn es einen Vektorunterraum $U \subset V$ und ein $v \in V$ gibt, so daß

$$A = v + U = \{v + u \mid u \in U\}.$$

U ist in diesem Fall eindeutig bestimmt und heißt *der Richtungsvektorraum zu* A. Man setzt

$$\dim A := \dim U$$

Die Lösungsmenge einer inhomogenen linearen Gleichung $f(x) = w$ ist also entweder leer, oder ein affiner Unterraum mit Richtungsvektorraum $\operatorname{Kern} f$.

8 Lineare Abbildungen und Matrizen

Wir haben gesehen, daß (m,n)-Matrizen über K lineare Abbildungen $K^m \to K^n$ induzieren. Wir wollen nun Matrizen zur Beschreibung linearer Abbildungen zwischen allgemeinen endlich erzeugten Vektorräumen benutzen.

Definition 24 (Darstellungsmatrix). Seien V und W Vektorräume mit Basen

$$\mathbf{v} = (v_1, \ldots, v_n) \text{ bzw. } \mathbf{w} = (w, \ldots, w_m).$$

Sei $f : V \to W$ linear. Dann ist $f(v_j)$ für jedes $j \in \{1, \ldots, n\}$ eine eindeutig bestimmte Linearkombination

$$f(v_j) = \sum_{i=1}^{m} a_{ij} w_i,$$

deren Koeffizienten a_{ij} natürlich auch von j abhängen! Damit erhält man eine (m,n)-Matrix

$$A^{\mathbf{v}}_{\mathbf{w}}(f) = (a_{ij}),$$

die *Darstellungsmatrix* oder *Koeffizientenmatrix* von f bezüglich der Basen $\mathbf{v}$ und $\mathbf{w}$.

Die j-te Spalte von $A^{\mathbf{v}}_{\mathbf{w}}(f)$ ist also gerade der Koordinatenvektor von $f(v_j)$ bezüglich der Basis $\mathbf{w}$.

Lemma 6. *Seien V und W Vektorräume mit Basen $\mathbf{v} = (v_1, \ldots, v_n)$ bzw. $\mathbf{w} = (w, \ldots, w_m)$. Mit $\Phi_{\mathbf{v}}, \Phi_{\mathbf{w}}$ bezeichnen wir die entsprechenden Koordinatensysteme, vgl. Definition 19. Sei $f : V \to W$ linear und sei A die Darstellungsmatix von f bezüglich der Basen $\mathbf{v}, \mathbf{w}$. Dann ist das folgende Diagramm kommutativ:*

$$\begin{array}{ccccc} & V & \xrightarrow[f]{} & W & \\ \Phi_{\mathbf{v}} & \uparrow & & \uparrow & \Phi_{\mathbf{w}} \\ & K^n & \xrightarrow[f_A]{} & K^m & \end{array}$$

Das heißt, es gilt

$$f \circ \Phi_{\mathbf{v}} = \Phi_{\mathbf{w}} \circ f_A.$$

Beweis. $f \circ \Phi_{\mathbf{v}}$ und $\Phi_{\mathbf{w}} \circ f_A$ sind lineare Abbildungen von K^n nach W. Wir müssen also nur zeigen, daß sie auf der kanonischen Basis $(e_1, \ldots, e_n)$ von K^n übereinstimmen. Aber es ist

$$f \circ \Phi_{\mathbf{v}}(e_j) \underset{(17)}{=} f(v_j),$$

$$\Phi_{\mathbf{w}} \circ f_A(e_j) = \Phi_{\mathbf{w}}(\begin{pmatrix} a_{1j} \\ \vdots \\ a_{mj} \end{pmatrix}) = \sum_{i=1}^{m} a_{ij} w_i = f(v_j).$$

□

Die Bedeutung dieses Lemmas ist die folgende: Die Theorie linearer Abbildungen für die Standardräume K^n ist einfach die Matrixtheorie. Die Theorie für endlich-dimensionale Vektorräume entspricht der „K^n-Theorie“ eineindeutig, *wenn* man Basen wählt. Das macht das Lemma auf sehr anschauliche Weise deutlich. Wegen der Basisabhängigkeit ist die Entsprechung aber nicht kanonisch.

Beispiel 44. Im Vektorraum $V = \mathrm{Pol}_3(\mathbb{R})$ der Polynome vom Grad ≤ 3 betrachten wir die Basis

$$\mathbf{v} = (1, x, x^2, x^3).$$

Im Vektorraum $W = \mathrm{Pol}_2(\mathbb{R})$ betrachten wir die beiden Basen

$$\mathbf{w_1} = (1, x, x^2)$$

und

$$\mathbf{w_2} = (1, x-1, (x-1)^2).$$

Wir bestimmen die Darstellungsmatrizen von

$$D : V \to W, f(x) \mapsto f'(x)$$

bezüglich der Basen $\mathbf{v}, \mathbf{w_1}$ und bezüglich der Basen $\mathbf{v}, \mathbf{w_2}$.

Wir erinnern: Die j-te Spalte von $A^{\mathbf{v}}_{\mathbf{w}}(f)$ ist gerade der Koordinatenvektor von $f(v_j)$ bezüglich der Basis $\mathbf{w}$.

Wegen $Dx^j = jx^{j-1}$ ist

$$A^{\mathbf{v}}_{\mathbf{w_1}}(D) = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \end{pmatrix}.$$

Benutzt man im Zielraum nun die Basis $\mathbf{w_2}$, so erhält man wegen

$$Dx^2 = 2x = 2 + 2(x-1), \quad Dx^3 = 3x^2 = 3 + 6(x-1) + 3(x-1)^2$$

die Darstellungsmatrix

$$A^{\mathbf{v}}_{\mathbf{w_2}}(D) = \begin{pmatrix} 0 & 1 & 2 & 3 \\ 0 & 0 & 2 & 6 \\ 0 & 0 & 0 & 3 \end{pmatrix}.$$

Schließlich betrachten wir noch die beiden Basen

$$(x, x^2, x^3, 1) \text{ für } V$$

und

$$(1, 2x, 3x^2) \text{ für } W.$$

Dann ergibt sich

$$A^{(x,x^2,x^3,1)}_{(1,2x,3x^2)}(D) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}.$$

In diesem Fall sind die Basen etwas „kraus“, dafür ist die Darstellungsmatrix denkbar einfach.

$\square$

Das vorstehende Beispiel demonstriert die Abhängigkeit der Darstellungsmatrix von den gewählten Basen. Daraus ergeben sich folgende Fragen: Gegeben $f : V \to W$ linear. Welche Matrizen können als Darstellungsmatrix von f auftreten? Gibt es darunter besonders einfache? Wie sehen die aus? Wir kommen auf diese Fragen später zurück.

Jetzt untersuchen wir die Frage, wie die Darstellungmatrix der Komposition von zwei linearen Abbildungen aussieht.

Satz 15. *Gegeben seien endlich-dimensionale K-Vektorräume U, V, W mit Basen*

$$\mathbf{u} = (u_1, \ldots, u_p) \text{ bzw. } \mathbf{v} = (v_1, \ldots, v_n) \text{ bzw. } \mathbf{w} = (w_1, \ldots, w_m).$$

Weiter seien

$$g : U \to V, \quad f : V \to W$$

lineare Abbildungen, die bezüglich der obigen Basen die Darstellungsmatrizen

$$A_{\mathbf{w}}^{\mathbf{v}}(f) = (a_{ij})$$

und

$$A_{\mathbf{v}}^{\mathbf{u}}(g) = (b_{jk})$$

haben. Dann ist die Darstellungsmatrix der Komposition $f \circ g : U \to W$

$$A_{\mathbf{w}}^{\mathbf{u}}(f \circ g) = (c_{ik})$$

gegeben durch

$$c_{ik} = \sum_{j=1}^{n} a_{ij} b_{jk}, \quad i \in \{1, \ldots, m\}, k \in \{1, \ldots, p\}. \tag{21}$$

Beweis.

$$f \circ g(u_k) = f\left(\sum_{j=1}^{n} b_{jk} v_j\right) = \sum_{j=1}^{n} b_{jk} f(v_j) = \sum_{j=1}^{n} b_{jk} \sum_{i=1}^{m} a_{ij} w_i = \sum_{i=1}^{m} \left(\sum_{j=1}^{n} a_{ij} b_{jk}\right) w_i.$$

□

Dieser Satz motiviert die folgende Definition:

Definition 25 (Matrixprodukt). Seien $A = (a_{ij})$ eine (p, n)-Matrix und $B = (b_{jk})$ eine (n, m)-Matrix. Wir definieren die (p, m)-Matrix $C = (c_{ik})$ durch (21) und nennen C das Produkt der Matrizen A und B:

$$AB := C.$$

Beachten Sie: Die Spaltenzahl der linken Matrix muß gleich der Zeilenzahl der rechten Matrix sein.

Damit kann man das Ergebnis des letzten Satzes auch so formulieren:

Satz 16. *Der Komposition linearer Abbildungen entspricht das Produkt ihrer Darstellungsmatrizen:*

$$A_{\mathbf{w}}^{\mathbf{u}}(f \circ g) = A_{\mathbf{w}}^{\mathbf{v}}(f) A_{\mathbf{v}}^{\mathbf{u}}(g).$$

Beispiel 45. Seien A eine (m, n)-Matrix und B eine (n, p)-Matrix. Diese induzieren lineare Abbildungen

$$f_A : K^n \to K^m \qquad f_B : K^p \to K^n,$$

deren Darstellungsmatrizen bezüglich der kanonischen Basen eben gerade A und B sind. Also findet man

$$\boxed{f_A \circ f_B = f_{AB}.}$$

□

Beispiel 46. Die Vektoren des K^n kann man als $(n, 1)$-Matrizen verstehen. Für eine (m, n)-Matrix A ist die Abbildung $f_A : K^n \to K^m$ dann einfach gegeben durch die Matrixmultiplikation

$$f_A(x) = Ax,$$

vgl. (19).

$\square$

Beispiel 47 (Koordinatenwechsel). Gegeben seien K-Vektorräume V, W mit jeweils zwei Basen

$$\mathbf{v} = (v_1, \ldots, v_n) \text{ bzw. } \tilde{\mathbf{v}} = (\tilde{v}_1, \ldots, \tilde{v}_n)$$

und

$$\mathbf{w} = (w_1, \ldots, w_m) \text{ bzw. } \tilde{\mathbf{w}} = (\tilde{w}_1, \ldots, \tilde{w}_m).$$

Weiter sei $f : V \to W$ linear. Wie unterscheiden sich die Darstellungsmatrizen von f bezüglich der Basen $\mathbf{v}, \mathbf{w}$ bzw. $\tilde{\mathbf{v}}, \tilde{\mathbf{w}}$? Es gilt

$$f = \mathrm{Id}_W \circ f \circ \mathrm{Id}_V$$

und deshalb

$$\boxed{A^{\tilde{\mathbf{v}}}_{\tilde{\mathbf{w}}}(f) = A^{\mathbf{w}}_{\tilde{\mathbf{w}}}(\mathrm{Id}) A^{\mathbf{v}}_{\mathbf{w}}(f) A^{\tilde{\mathbf{v}}}_{\mathbf{v}}(\mathrm{Id})}$$

Die Matrix $A^{\tilde{\mathbf{v}}}_{\mathbf{v}}(\mathrm{Id}) =: (c_{ij})$ heißt auch die Basistransformationsmatrix von der Basis $\mathbf{v}$ auf die Basis $\tilde{\mathbf{v}}$. Sie ist gegeben durch

$$v_j = \sum_j c_{ij} \tilde{v}_i.$$

Die Darstellungsmatizen von f bezüglich zweier Basispaare unterscheiden sich also durch Links- und Rechtsmultiplikation mit Basistransformationsmatrizen.

$\square$

9 Automorphismen und die allgemeine lineare Gruppe

Sei V ein endlich-dimensionaler K-Vektorraum.

Definition 26. Die Isomorphismen von V auf sich bilden bezüglich der Komposition eine Gruppe, die sogenannte *Automorphismengruppe*

$$\mathrm{Aut}(V).$$

Das neutrale Element ist die identische Abbildung

$$\mathrm{Id} : V \to V, v \mapsto v,$$

und das inverse Element ist die Unkehrabbildung $f^{-1} : V \to V$.

Bemerkung. Die Gruppe $\mathrm{Aut}(V)$ ist im allgemeinen nicht kommutativ. Dennoch folgt aus $f \circ g = \mathrm{Id}$ oder $g \circ f = \mathrm{Id}$ jeweils die andere Gleichung und es ist $g = f^{-1}$. Das kann man ganz allgemein aus den Gruppenaxiomen folgern. In unserem konkreten Fall ergibt es sich aber auch so: Aus $f \circ g = \mathrm{Id}$ folgt, daß f surjektiv ist. (Wie denn?) Dann ist es aber nach der Dimensionsformel auch injektiv, also besitzt es eine Umkehrabbildung f^{-1}, und aus $f \circ g = \mathrm{Id}$ folgt durch Anwendung von f^{-1}, daß $g = f^{-1}$. Dann ist aber auch $g \circ f = \mathrm{Id}$.

Beispiel 48. Wir betrachten den Fall $V = K^n$. Wir bezeichnen mit E oder gelegentlich auch E_n die (n,n)-Einheitsmatrix:

$$E = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{pmatrix}.$$

Mit dem sogenannten Kronecker-Symbol

$$\delta_{ij} := \begin{cases} 1 & \text{für } i = j \\ 0 & \text{sonst} \end{cases}$$

ist also $E = (\delta_{ij})$. Dann liefert

$$A \mapsto f_A$$

eine Bijektion zwischen den (n,n)-Matrizen und den linearen Abbildungen von V in sich. Es gilt $f_E = \mathrm{Id}$. Also folgt aus $f_A \circ f_B = f_{AB}$, daß

$$f_A \circ f_B = \mathrm{Id} \iff AB = E.$$

Nach der obigen Bemerkung ist das auch äquivalent zu $f_B \circ f_A = \mathrm{Id}$, also zu

$$BA = E.$$

Den Automorphismen von K^n entsprechen also eineindeutig die (n,n)-Matrizen A, zu denen es ein B mit $AB = E$ gibt.

□

Definition 27 (Invertierbare Matrizen). Man nennt eine (n,n)-Matrix A *invertierbar*, wenn es eine (n,n)-Matrix B gibt, so daß $AB = BA = E$. man schreibt dann $B = A^{-1}$. Die invertierbaren (n,n)-Matrizen bilden bezüglich der Matrixmultiplikation eine Gruppe mit E als neutralem Element, die sogenannte *Allgemeine Lineare Gruppe* $\mathrm{GL}(n)$.

Die Gruppeneigenschaften von $\mathrm{GL}(n)$ folgen unmittelbar aus denen von $\mathrm{Aut}(K^n)$, weil $A \mapsto f_A$ die Multiplikation erhält: Man nennt die beiden Gruppen auch *isomorph*; aber das ist der Isomorphie-Begriff für Gruppen, nicht der für Vektorräume.

Bemerkung. Sind $f : V \to W$ und $g : W \to V$ Abbildungen zwischen Mengen und gilt $f \circ g = \mathrm{Id}_W$, so folgt, daß f surjektiv und g injektiv ist. Ist aber $V = W$ endlich-dimensional und sind f und g linear, so folgt, daß f bijektiv und $g = f^{-1}$ ist. Entsprechend folgt für quadratische Matrizen aus einer der beiden Gleichungen $AB = E$ oder $BA = E$ die jeweils andere, und A und B sind invers zueinander. Dies kann man auch aus den Gruppenaxiomen beweisen.

Lemma 7. *Die (n,n)-Matrix A ist genau dann invertierbar, wenn*

$$\mathrm{Rang}\, A = n.$$

Beweis. Die Matrix A genau dann invertierbar, wenn $f_A : K^n \to K^n$ ein Isomorphismus ist. Das ist genau dann der Fall, wenn f_A surjektiv, d.h. wenn $\mathrm{Rang}\, f_A = n$ ist. Aber $\mathrm{Rang}\, f_A = \mathrm{Rang}\, A$. □

Beispiel 49 (Berechnung der Inversen Matrix). Um die Inverse der (n,n)-Matrix A zu finden, muß man die Gleichung

$$AX = E$$

lösen. Das ist nur möglich, wenn A invertierbar, d.h. wenn $\operatorname{Rang} A = n$ In diesem Fall liefert der Gaußalgorithmus aus A eine Dreiecksmatrix mit lauter von Null verschiedenen Diagonalelementen:

$$\begin{pmatrix} a'_{11} & \dots & a'_{1n} \\ & \ddots & \vdots \\ 0 & & a'_{nn} \end{pmatrix}.$$

Wendet man darauf noch einmal den Gaußalgorithmus von unten und hinten beginnend an, so ergibt sich eine Diagonalmatrix mit denselben Diagonalelementen:

$$\begin{pmatrix} a'_{11} & & 0 \\ & \ddots & \\ 0 & & a'_{nn} \end{pmatrix}.$$

Wenden wir dieselben Operationen (Gauß runter und rauf) auf die rechts um den i-ten Einheitsvektor $e_j = (0, \dots, 1, \dots, 0)^T$ erweiterte Matrix $(A|e_j)$ an und teilen schließlich die i-te Zeile noch durch a'_{ii}, so steht in der rechten Spalte die Lösung der Gleichung $Ax = e_j$. Wir bezeichnen sie mit x_j. Man kann diese Prozedur auch gleichzeitig für alle e_j durchführen: Man erweitert die Matrix A rechts um die Einheitsmatrix E, deren Spalten ja gerade die Vektoren e_j sind, und wendet dann dieselben Operationen auf die so erweiterte Matrix $(A|E)$ an. In der j-ten Spalte rechts steht dann x_j. Bezeichnet man die sich rechts ergebende Matrix mit $X = (x_1, \dots, x_n)$, so ist nach Definition der Matrixmultiplikation $AX = E$.

Also ist bei Rang $A = n$ die Gleichung lösbar, sogar eindeutig lösbar, und wir haben einen **Algorithmus** für ihre Lösung angegeben. Ist die Gleichung nicht lösbar, so ergibt sich das auch bei dem Algorithmus: die Zeilenstufenform von A hat dann Nullzeilen.

$\square$

Beispiel 50. Sei $A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$. Das oben beschriebene Verfahren liefert:

$$\begin{pmatrix} 1 & 2 & | & 1 & 0 \\ 3 & 4 & | & 0 & 1 \end{pmatrix} \to \begin{pmatrix} 1 & 2 & | & 1 & 0 \\ 0 & -2 & | & -3 & 1 \end{pmatrix} \to \begin{pmatrix} 1 & 0 & | & -2 & 1 \\ 0 & -2 & | & -3 & 1 \end{pmatrix} \to \begin{pmatrix} 1 & 0 & | & -2 & 1 \\ 0 & 1 & | & \frac{3}{2} & -\frac{1}{2} \end{pmatrix}.$$

Also ist $A^{-1} = \begin{pmatrix} -2 & 1 \\ \frac{3}{2} & -\frac{1}{2} \end{pmatrix}$.

$\square$

Beispiel 51. Sei $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, und sei $\Delta := ad - bc \neq 0$. Wir nehmen der Einfachheit halber noch an, daß $a \neq 0$; das Ergebnis ist davon allerdings unabhängig. Das oben beschriebene Verfahren liefert:

$$\begin{pmatrix} a & b & | & 1 & 0 \\ c & d & | & 0 & 1 \end{pmatrix} \to \begin{pmatrix} a & b & | & 1 & 0 \\ 0 & d - \frac{c}{a}b & | & -\frac{c}{a} & 1 \end{pmatrix} \to \begin{pmatrix} a & 0 & | & \frac{ad}{\Delta} & \frac{ab}{\Delta} \\ 0 & \frac{\Delta}{a} & | & -\frac{c}{a} & 1 \end{pmatrix}$$
$$\to \begin{pmatrix} 1 & 0 & | & \frac{d}{\Delta} & -\frac{b}{\Delta} \\ 0 & 1 & | & -\frac{c}{\Delta} & \frac{a}{\Delta} \end{pmatrix}.$$

Also ist $A^{-1} = \frac{1}{\Delta} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$.

$\square$

10 Vektorräume von linearen Abbildungen und Matrizen

Die Menge der Abbildungen W^M von einer beliebigen Menge $M \neq \emptyset$ in einen K-Vektorraum W bildet mit wertweiser Addition und Skalarmultiplikation wieder einen K-Vektorraum. Das ist insbesondere der Fall, wenn $M = V$ selbst ein K-Vektorraum ist. Die Menge der *linearen* Abbildungen von V nach W bildet dann einen Vektorunterraum.

Definition 28. Seien V, W zwei K-Vektorräume. Wir bezeichen mit

$$\mathrm{Hom}(V, W) := \{f : V \to W \mid f \text{ linear}\}$$

den K-Vektorraum der linearen Abbildungen von V nach W. Wir setzen

$$\mathrm{End}(V) := \mathrm{Hom}(V, V).$$

Definition 29. Für positive natürliche Zahlen $m, n \in \mathbb{N}$ bezeichnen wir mit

$$M(m, n; K)$$

die Menge der (m, n)-Matrizen mit Koeffizienten in K. Mit der koeffizientenweisen Addition und Skalarmultiplikation

$$\begin{aligned}(a_{ij}) + (b_{ij}) &:= (a_{ij} + b_{ij}),\\ \lambda(a_{ij}) &:= (\lambda a_{ij})\end{aligned}$$

wird $M(m, n; K)$ ein K-Vektorraum. Wir verallgemeinern nun den Satz 14:

Satz 17. *Seien V, W zwei endlich-dimensionale K-Vektorräume mit Basen*

$$\mathbf{v} = (v_1, \ldots, v_n), \quad \mathbf{w} = (w_1, \ldots, w_m).$$

Dann ist die Abbildung

$$A_{\mathbf{w}}^{\mathbf{v}} : \mathrm{Hom}(V, W) \to M(m, n; K)$$

ein Isomorphismus.

Beweis. Zu zeigen wäre, daß die Abbildung linear ist. Das ist sehr einfach, ich mache es nicht vor. Die Abbildung ist bijektiv, weil es nach dem Existenz- und Eindeutigkeitssatz für lineare Abbildungen zu einer gegebenen Matrix (a_{ij}) genau eine lineare Abbildung $f : V \to W$ gibt, für die

$$f(v_j) = \sum a_{ij} w_i$$

ist. □

Offenbar bilden die Matrizen E_{ij}, die in der j-ten Spalte der i-ten Zeile eine 1 und sonst lauter Nullen enthalten, eine Basis von $M(m, n; K)$, und deshalb ist

$$\boxed{\dim M(m, n; K) = mn.}$$

Korollar 4. *Seien V, W zwei K-Vektorräume endlicher Dimension. Dann gilt*

$$\boxed{\dim \mathrm{Hom}(V, W) = \dim V \; \dim W.}$$

Bemerkung: Die Algebra der Endomorphismen/quadratischen Matrizen. Die Addition von $\mathrm{Hom}(V, W)$ ist mit der Komposition „verträglich“: Es gilt

$$(f+g)\circ h = f\circ h + g\circ h, \quad f\circ(g+h) = f\circ g + f\circ h,$$

wenn immer diese Operationen sinnvoll sind, insbesondere, wenn $f, g, h \in \mathrm{End}(V)$ sind. Damit hat man in $\mathrm{End}(V)$ also eine Addition und eine damit verträgliche Multiplikation. Man sagt $\mathrm{End}(V)$ ist ein *Ring* mit Id als Einselement. Weil (wenigstens für $V \neq K$) nicht alle $f \in \mathrm{End}(V)\backslash\{0\}$ invertierbar sind, ist $\mathrm{End}(V)$ kein Körper.

Andrerseits gibt es zusätzlich eine Skalarmultiplikation, die $\mathrm{End}(V)$ zum K-Vektorraum macht und die mit der obigen Multiplikation verträglich ist:

$$(\lambda f)\circ g = f\circ(\lambda g) = \lambda(f\circ g),$$

Einen Ring, der gleichzeitig ein K-Vektorraum ist und dessen beide „Multiplikationen“ sich auf diese Weise vertragen, nennt man eine *K-Algebra*.

Wählt man $V = K^n, W = K^m$ so übertragen sich diese Bemerkungen unmittelbar auf $M(n, n; K)$.

Einen besonders einfachen Vektorraum linearer Abbildungen erhält man, wenn $W = K^1 = K$ ist.

Definition 30 (Dualraum). Sei V ein endlich-dimensionaler K-Vektorraum. Dann nennt man

$$V^* := \mathrm{Hom}(V, K)$$

den *Dualraum* von V. Die Elemente von V^* nennt man auch *Linearformen* auf V. Ist

$$\mathbf{v} = (v_1, \ldots, v_n)$$

eine Basis von V, so gibt es zu jedem $i \in \{1, \ldots, n\}$ genau eine lineare Abbildung $v_i^* : V \to K$ mit

$$v_i^*(v_j) = \begin{cases} 1 & \text{für } i = j \\ 0 & \text{für } i \neq j. \end{cases}$$

Dann ist $\mathbf{v}^* = (v_1^*, \ldots, v_n^*)$ eine Basis von V^*, die sogenannnte *duale Basis* zur Basis $\mathbf{v}$.

Beweis der Basiseigenschaft. Wir wissen schon, daß

$$\dim V^* = n \cdot 1 = n.$$

Daher genügt der Nachweis, daß $(v_1^*, \ldots, v_n^*)$ linear unabhängig ist. Aus

$$\sum_{i=1}^{n} \lambda_i v_i^* = 0$$

folgt aber insbesondere für jedes $j \in \{1, \ldots, n\}$

$$0 = \sum_{i=1}^{n} \lambda_i v_i^*(v_j) = \lambda_j.$$

□

Bemerkung. Ist V ein Vektorraum mit Basis $\mathbf{v} = (v_1, \ldots, v_n)$ und dualer Basis $(v_1^*, \ldots, v_n^*)$, so gilt für alle $v \in V$

$$\boxed{v = \sum_{i=1}^{n} v_i^*(v) v_i.}$$

Die duale Basis liefert also gerade die Koeffizienten von v in der Entwicklung nach der Basis $\mathbf{v}$.

Definition 31 (Duale Abbildung). Sei $f : V \to W$ eine lineare Abbildung endlich-dimensionaler Vektorräume. Dann gibt es genau eine lineare Abbildung

$$f^* : W^* \to V^*$$

mit

$$f^*(\omega)(v) = \omega(f(v)) \tag{22}$$

für alle $v \in V$ und $\omega \in W^*$. Diese Abbildung heißt *die duale Abbildung* zu f.

Beweis. Die Eindeutigkeit ist klar, weil $f^*(\omega)$ durch (22) vollständig beschrieben ist.

Wir zeigen die Existenz: Sei $\omega \in W^*$. Dann ist

$$v \mapsto \omega(f(v))$$

als Komposition von zwei linearen Abbildungen linear, also ein Element von V^*. Damit definiert (22) also eine Abbildung

$$f^* : W^* \to V^*.$$

Wir müssen noch zeigen, daß f^* linear ist. Aber

$$\begin{aligned} f^*(\omega_1 + \omega_2)(v) &= (\omega_1 + \omega_2)(f(v)) = \omega_1(f(v)) + \omega_2(f(v)) \\ &= f^*(\omega_1)(v) + f^*(\omega_2)(v). \end{aligned}$$

Ebenso folgt $f^*(\lambda\omega) = \lambda f^*(\omega)$. □

Für den folgenden Satz brauchen wir eine

Definition 32 (Transponierte Matrix). Sei

$$A = (a_{ij})_{i=1,\ldots,m;\, j=1,\ldots,n}$$

eine (m,n)-Matrix. Durch Vertauschen von Zeilen und Spalten erhält man daraus eine (n,m)-Matrix

$$A^T = (a^T_{ij})_{i=1,\ldots,n;\, j=1,\ldots,m}.$$

Diese heißt die *Transponierte* von A. Es ist also

$$a^T_{ij} = a_{ji}, \quad i \in \{1,\ldots,n\}, j \in \{1,\ldots,m\}.$$

Satz 18 (Duale Abbildung). *Seien V, W zwei endlich-dimensionale K-Vektorräume mit Basen $\mathbf{v}$ bzw. $\mathbf{w}$. Seien $f : V \to W$ eine lineare Abbildung und $f^* : W^* \to V^*$ die duale Abbildung. Dann gilt:*

$$\operatorname{Rang} f = \operatorname{Rang} f^*. \tag{23}$$

Sind weiter $\mathbf{v}^*$ und $\mathbf{w}^*$ die dualen Basen, so gilt

$$A^{\mathbf{w}^*}_{\mathbf{v}^*}(f^*) = \left(A^{\mathbf{v}}_{\mathbf{w}}(f)\right)^T. \tag{24}$$

Beweis. Zu (23). Sei $n := \dim V$ und $r = \operatorname{Rang} f$. Wie im Beweis der Dimensionsformel findet man eine Basis $v_1, \ldots, v_n$ von V, so daß $v_{r+1}, \ldots, v_n$ eine Basis von $\operatorname{Kern} f$ und $w_1 := f(v_1), \ldots, w_r = f(v_r)$ eine Basis von $\operatorname{Bild} f$ ist. Die letztere ergänzen wir durch $w_{r+1}, \ldots, w_m$ zu einer Basis von W. Wir bezeichnen mit $\mathbf{v}^*$ und $\mathbf{w}^*$ die dualen Basen. Dann gilt

$$f^*(w_i^*)(v_j) = \begin{cases} w_i^* f(v_j)) = 0 & \text{für } j > r \\ w_i^*(w_j) = v_i^*(v_j) & \text{für } j \leq r. \end{cases}$$

Daher ist

$$f^*(w_i^*) = 0 \text{ für alle } i > r$$

und

$$f^*(w_i^*) = v_i^* \text{ für alle } i \leq r.$$

Daraus folgt $\operatorname{Rang} f^* = r = \operatorname{Rang} f$.

Zu (24). Seien jetzt $\mathbf{v}, \mathbf{w}$ beliebige Basen und $\mathbf{v}^*$ und $\mathbf{w}^*$ die dualen Base dazu. Dann ist

$$\begin{aligned} f^*(w_i^*)(v_k) = w_i^*(f(v_k)) = w_i^*(\sum_{j=1}^m a_{jk} w_j) &= \sum_{j=1}^m a_{jk} w_i^*(w_j) = a_{ik} \\ = a_{ki}^T = \sum_{j=1}^n a_{ji}^T v_j^*(v_k). \end{aligned}$$

Also ist

$$f^*(w_i^*) = \sum_{j=1}^n a_{ji}^T v_j^*.$$

□

Als Korollar ergibt sich der folgende Satz über Matrizen

Satz 19 (Zeilenrang=Spaltenrang). *Sei $A = (a_{ij})$ eine (m,n)-Matrix. Dann sind die Spalten von A Vektoren im K^m. Die Dimension des von diesen aufgespannten Unterraums des K^m heißt der* Spaltenrang *von A. Er ist nach dem Basissatz gleich der Maximalzahl linear unabhängiger Spalten von A.*
Entsprechend kann man die Zeilen (senkrecht geschrieben) als Vektoren im K^n interpretieren. Die Dimension des von ihnen aufgespannten Unterraums des K^n heißt der Zeilenrang *von A. Er ist nach dem Basissatz gleich der Maximalzahl linear unabhängiger Zeilen von A. Es gilt*

$$\text{Zeilenrang von } A = \text{Spaltenrang von } A = \operatorname{Rang} f_A.$$

Diese Zahl nennt man auch den Rang der Matrix *A und bezeichnet sie mit*

$$\operatorname{Rang} A.$$

Für den Beweis benötigen wir noch ein kleines Lemma. Wir hatten den Rang einer Matrix A definiert als den Rang von f_A. Allgemeiner folgt aus dem Lemma 6 sofort:

Lemma 8. *Ist $f : V \to W$ linear und haben V, W endliche Basen $\mathbf{v}$ bzw. $\mathbf{w}$, so ist*

$$\operatorname{Rang} A_{\mathbf{w}}^{\mathbf{v}}(f) = \operatorname{Rang} f.$$

Beweis des Korollars. Die Spaltenvektoren von A sind gerade die Bilder der Basisvektoren e_i unter der linearen Abbildung $f_A : K^n \to K^m$. Daher ist

$$\text{Spaltenrang von } A = \operatorname{Rang} f_A.$$

Der Zeilenrang von A ist gerade der Spaltenrang der transponierten Matrix A^T, und die ist die Darstellungsmatrix der dualen Abbildung $(f_A)^* : (K^m)^* \to (K^n)^*$. Also ist nach dem Lemma

$$\text{Zeilenrang von } A = \operatorname{Rang}(f_A)^*.$$

Nach dem letzten Satz ist aber $\operatorname{Rang}(f_A)^* = \operatorname{Rang} f_A$. □

Alternative Beweisidee. Der Rang eines linearen Gleichungssystems mit Systemmatrix A war n – Dimension des Lösungsraumes, also

$$n - \dim \operatorname{Kern} f_A = \operatorname{Rang} f_A = \operatorname{Rang} A.$$

Also kann man den Rang einer Matrix A mit dem Gaußalgorithmus bestimmen: Die Anzahl der verbleibenden nicht-trivialen Zeilen ist $\operatorname{Rang} A$. Man kann nun leicht überlegen, daß die Zeilenoperationen des auf eine Matrix angewendeten Gaußalgorithmus den Spann der Zeilen und damit die Dimension des Zeilenraums nicht verändern. In Zeilenstufenform ist aber die lineare Unabhängigkeit der nicht-trivialen Zeilen und damit der Zeilenrang offensichtlich. □

Bemerkung (Unabhängigkeitstest im K^n). Um die lineare Unabhängigkeit von k Vektoren des K^n zu testen, schreibt man diese Vektoren als Zeilen *oder Spalten* in eine Matrix und bestimmt deren Rang mit dem Gaußalgorithmus.

11 Euklidische Vektorräume

Wir betrachten nun Vektorräume, in denen wir eine Längen- und Winkelmessung gegeben haben. Dazu braucht man eine zusätzliche Struktur, ein sogenanntes Skalarprodukt. Der Körper ist in diesem Abschnitt stets der Körper $\mathbb{R}$ der reellen Zahlen:

$$\boxed{K = \mathbb{R}.}$$

Definition 33 (Skalarprodukt). Sei V ein $\mathbb{R}$-Vektorraum.

1. Eine *Bilinearform auf* V ist eine Abbildung

$$\langle ., . \rangle : V \times V \to \mathbb{R}, (v, w) \mapsto \langle v, w \rangle$$

   mit folgenden Eigenschaften:

   (a) Für alle $v_1, v_2, w \in V$ und $\lambda_1, \lambda_2 \in \mathbb{R}$ gilt

$$\langle \lambda_1 v_1 + \lambda_2 v_2, w \rangle = \lambda_1 \langle v_1, y \rangle + \lambda_2 \langle v_2, w \rangle.$$

   (b) Für alle $v, w_1, w_2 \in V$ und $\lambda_1, \lambda_2 \in \mathbb{R}$ gilt

$$\langle v, \lambda_1 w_1 + \lambda_2 w_2 \rangle = \lambda_1 \langle v, w_1 \rangle + \lambda_2 \langle v, w_2 \rangle.$$

2. Eine Bilinearform $\langle ., . \rangle$ auf V heißt *symmetrisch*, wenn für alle $v, w \in V$

$$\langle v, w \rangle = \langle w, v \rangle.$$

3. Eine symmetrische Bilinearform $\langle ., . \rangle$ auf V heißt ein *Skalarprodukt*, wenn sie überdies *positiv definit* ist, d.h. wenn für alle $v \in V \backslash \{0\}$

$$\langle v, v \rangle \; > \; 0.$$

4. Ein Paar $(V, \langle ., . \rangle)$ aus einem <u>endlich-dimensionalen</u> $\mathbb{R}$-Vektorraum V mit einem Skalarprodukt $\langle ., . \rangle$ heißt ein *euklidischer Vektorraum.*

5. Ist $\langle ., . \rangle$ ein Skalarprodukt auf V, so definieren wir die dadurch induzierte *Norm* durch

$$\|v\| := \sqrt{\langle v, v \rangle} \text{ für alle } v \in V.$$

   Wir nennen $\|v\|$ auch die *Länge* des Vektors $\|v\|$.

Beispiel 52 (Standardprodukt auf dem $\mathbb{R}^n$). Für $x = (x_1, \ldots, x_n), y = (y_1, \ldots, y_n) \in \mathbb{R}^n$ definiere

$$\langle x, y \rangle := \sum_{i=0}^{n} x_i y_i.$$

Das ist ein Skalarprodukt, das sogenannte kanonische Skalarprodukt auf $\mathbb{R}^n$. Die zugehörige Norm ist

$$\|x\| = \sqrt{\sum_{i=1}^{n} x_i^2}.$$

$\square$

Beispiel 53. Auf $C^0([0,1])$ definiert

$$\langle f,g\rangle := \int_0^1 f(x)g(x)dx$$

ein Skalarprodukt.

□

Definition 34 (Norm). Eine *Norm* auf einem $\mathbb{R}$-Vektorraum V ist eine Abbildung

$$\|.\| : V \to \mathbb{R}$$

mit folgenden Eigenschaften:

1. Für alle $x \in V$ ist
$$\|v\| \geq 0$$
und Gleichheit steht nur für $v = 0$.
2. Für alle $\lambda \in \mathbb{R}$ und $v \in V$ gilt
$$\|\lambda v\| = |\lambda|\,\|v\|.$$
3. Für alle $v,w \in V$ gilt die „Dreiecksungleichung“
$$\|v+w\| = \|v\ +\ \|w\|.$$

Bemerkung: Die von einem Skalarprodukt induzierte Norm erfüllt offenbar die beiden ersten Axiome, die Dreiecksungleichung zeigen wir im übernächsten Satz. Damit wir dann klar, daß die von einem Skalarprodukt induzierte Norm wirklich eine Norm im Sinne obiger Definition ist. Zunächst brauchen wir aber den folgenden

Satz 20 (Cauchy-Schwarzsche Ungleichung). *Sei $\langle .,.\rangle$ ein Skalarprodukt auf den reellen Vektorraum V. Dann gilt für alle $v,w \in V$ die sogenannte* Cauchy-Schwarzsche Ungleichung

$$|\langle v,w\rangle| \leq \|v\|\,\|w\|.$$

Beweis. Zunächst ist für jede relle Zahl λ

$$0 \leq \langle v+\lambda w, v+\lambda w\rangle = \langle v,v\rangle + \lambda^2\langle w,w\rangle + 2\lambda\langle v,w\rangle.$$

Wenn man λ so wählt, daß die rechte Seite möglichst klein wird, ist diese Aussage am stärksten. Das Minimum der rechten Seite in Abhängigkeit von λ können Sie aber einfach durch Differenzieren bestimmen (Ableitung = 0 setzen!) und erhalten $\lambda = -\frac{\langle v,w\rangle}{\langle w,w\rangle}$.

Wir können dabei ruhig voraussetzen, daß $w \neq 0$, denn sonst ist die Behauptung sowieso klar.

Setzen wir das so gefundene λ oben ein, so folgt

$$0 \leq \langle v,v\rangle + \frac{\langle v,w\rangle^2}{\langle w,w\rangle^2}\langle w,w\rangle - 2\frac{\langle v,w\rangle^2}{\langle w,w\rangle} = \langle v,v\rangle - \frac{\langle v,w\rangle^2}{\langle w,w\rangle}.$$

Daraus folgt nach Multiplikation mit $\langle w,w\rangle$, daß $0 \leq \langle v,v\rangle\langle w,w\rangle - \langle v,w\rangle^2$, und das war die Behauptung. □

Satz 21 (Dreiecksungleichung). *Ist* $\|.\|$ *die vom Skalarprodukt* $\langle .,. \rangle$ *auf* V *induzierte Norm, so gilt für alle* $v, w \in V$

$$\|v + w\| \ \leq \ \|v\| + \|w\|.$$

Beweis. Es ist für $v, w \in V$

$$(\|v\| + \|w\|)^2 = \|v\|^2 + 2\|v\| \, \|w\| + \|w\|^2 \geq \|v\|^2 + 2\langle v, w\rangle + \|w\|^2 =< v + w, v + w >$$

Also folgt

$$\|v\| + \|w\| \geq \|v + w\|.$$

□

Damit können wir nun unter Verwendung von Eigenschaften der Cosinusfunktion den Winkel definieren:

Definition 35 (Winkel). Der Winkel zwischen den Vektoren $v, w \in V\backslash\{0\}$ eines $\mathbb{R}$-Vektorraumes V mit Skalarprodukt ist die eindeutig bestimmte Zahl $\phi \in [0, \pi]$ mit

$$\cos\phi = \frac{\langle v, w\rangle}{\|v\| \, \|w\|}.$$

Beachten Sie, daß die rechte Seite nach der Cauchy-Schwarzschen Ungleichung zwischen -1 und 1 liegt.

Satz 22 (Polarisation). *In einem euklidischen Vektorraum gilt für alle* v, w

$$\langle v, w\rangle = \frac{1}{4}\left(\|v + w\|^2 - \|v - w\|^2\right).$$

Das Skalarprodukt ist also durch die zugehörige Norm eindeutig bestimmt.

Beweis. Es gilt

$$\begin{aligned}\|v + w\|^2 - \|v - w\|^2 &=< v + w, v + w > - < v - w, v - w > \\ &= \langle v, v\rangle + \langle w, w\rangle + 2\langle v, w\rangle - (\langle v, v\rangle + \langle w, w\rangle - 2\langle v, w\rangle) \\ &= 4\langle v, w\rangle.\end{aligned}$$

□

Satz 23 (Parallelogrammgleichung). *In einem euklidischen Vektorraum gilt für alle* v, w

$$2(\|v\|^2 + \|w\|^2) = \|v + w\|^2 + \|v - w\|^2.$$

Die Quadratsumme der vier Seiten eines Parallelogramms ist gleich der Quadratsumme der beiden Diagonalen.

Beweis. Analog zum vorstehenden Satz. □

Definition 36 (Orthogonalität). Sei $(V, \langle .,. \rangle)$ ein $\mathbb{R}$-Vektorraum mit Skalarprodukt.

1. $v_1, \ldots, v_k \in V$ heißen *orthogonal*, wenn

$$\langle v_i, v_j \rangle = 0 \text{ für alle } i \neq j.$$

2. $v_1, \ldots, v_k \in V$ heißen *orthonormal* oder *ein Orthonormalsystem*, wenn sie orthogonal sind und außerdem

$$\|v_i\| = 1 \text{ für alle } i \in \{1, \ldots, n\}.$$

   Vektoren v mit $\|v\| = 1$ heißen auch *Einheitsvektoren*. Die Orthonormalitäts-Bedingung kann man auch schreiben als

$$\langle v_i, v_j \rangle = \delta_{ij}$$

   mit dem sogenannten Kronecker-Delta

$$\delta_{ij} = \begin{cases} 1 & \text{für } i = j \\ 0 & \text{sonst.} \end{cases}$$

Lemma 9 (Fourierkoeffizienten). *Sind $v_1, \ldots, v_k \in V$ orthonormal, und ist*

$$v = \sum_{i=1}^{k} \lambda_i v_i,$$

so folgt für alle $j \in \{1, \ldots, k\}$

$$\lambda_j = \langle v, v_j \rangle.$$

Insbesondere folgt (mit $v = 0$) die lineare Unabhängigkeit von $(v_1, \ldots, v_k)$.

Beweis. Aus

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

folgt durch Skalarmultiplikation mit v_j

$$\langle v, v_j \rangle = \langle \sum_{i=1}^{k} \lambda_i v_i, v_j \rangle = \sum_{i=1}^{k} \lambda_i \langle v_i, v_j \rangle = \lambda_j.$$

□

Beispiel 54 (Trigonometrische Fourierkoeffizienten). Auf $C^0([0, 2\pi])$ definiert

$$\langle f, g \rangle := \frac{1}{\pi} \int_0^{2\pi} f(x) g(x) dx$$

ein Skalarprodukt. In der Analysis zeigt man, daß bezüglich dieses Skalarproduktes die Funktionen

$$1/\sqrt{2}, \sin x, \cos x, \sin 2x, \cos 2x, \sin 3x, \cos 3x, \ldots$$

orthonormal sind. Für $f \in C^0([0, 2\pi])$ heißen dann

$$a_k := \langle f, \cos kx \rangle = \frac{1}{\pi} \int_0^{2\pi} f(x) \cos kx dx,$$
$$b_k := \langle f, \sin kx \rangle = \frac{1}{\pi} \int_0^{2\pi} f(x) \sin kx dx$$

die Fourierkoeffizienten von f. In der Fourieranalyse stellt man f dann dar als (unendliche) Linearkombination der Sinus/Cosinus-Funktionen. Die Koeffizienten dieser Darstellung sind gerade die a_k und b_k:

$$f(x) \sim \frac{a_0}{2} + \sum_k (a_k \cos kx + b_k \sin kx).$$

□

Satz 24 (Orthonormalisierungsverfahren von Gram/Schmidt). *Seien* $(V, \langle ., . \rangle)$ *ein Vektorraum mit Skalarprodukt und* $w_1, \ldots, w_k \in V$ *linear unabhängig. Dann gibt es eindeutig bestimmte*

$$\lambda_{ij} \in \mathbb{R}, \quad 1 \leq i \leq j \leq k,$$

mit folgenden Eigenschaften:

1. $\lambda_{ii} > 0$ *für alle* $i \in \{1, \ldots, k\}$.
2. *Die durch*
$$v_j := \sum_{i=1}^{j} \lambda_{ij} w_i$$
definierten Vektoren $v_1, \ldots, v_k$ *sind orthonormal.*

Insbesondere ist $\mathrm{Spann}(v_1, \ldots, v_j) \subset \mathrm{Spann}(w_1, \ldots, w_j)$, *und aus Dimensionsgründen folgt*

$$\mathrm{Spann}(v_1, \ldots, v_j) = \mathrm{Spann}(w_1, \ldots, w_j) \quad \textit{für alle } j \in \{1, \ldots, k\}.$$

Die $v_1, \ldots, v_k$ *kann man durch ein rekursives Verfahren aus den* w_i *gewinnen, das sogenannte Orthonormalisierungsverfahren von Gram und Schmidt, siehe die Gleichungen* (27) *und* (28) *im nachstehenden Beweis.*

Beweis. Der Beweis erfolgt durch vollständige Induktion über k.

$\underline{k = 1.}$ Dann haben $\lambda_{11} = \frac{1}{\|w_1\|}$ und $v_1 = \lambda_{11} w_1$ die gewünschten Eigenschaften. λ_{11} ist offensichtlich eindeutig bestimmt.

$\underline{k \implies k+1.}$ Seien $w_1, \ldots, w_{k+1}$ linear unabhängig. Nach Induktionsvoraussetzung gibt es eindeutig bestimmt $\lambda_{ij}, 1 \leq i \leq j \leq k$ mit $\lambda_{ii} > 0$, so daß die damit bestimmten $v_1, \ldots, v_k$ orthonormal sind.

Wir zeigen zunächst die Einzigkeit der $\lambda_{i\,k+1}$. Sei also

$$v_{k+1} = \sum_{i=1}^{k+1} \lambda_{i\,k+1} w_i = \lambda_{k+1\,k+1} w_{k+1} + \underbrace{\sum_{i=1}^{k} \lambda_{i\,k+1} w_i}_{\in \mathrm{Spann}(v_1, \ldots, v_k)}$$

orthogonal zu $v_1, \ldots, v_k$, von der Länge 1 und mit $\lambda_{k+1\,k+1} > 0$. Dann gibt es $\mu_{i\,k+1}$ mit

$$v_{k+1} = \lambda_{k+1\,k+1} w_{k+1} + \sum_{i=1}^{k} \mu_{i\,k+1} v_i \tag{25}$$

Multiplikation mit $v_j, 1 \leq j \leq k$ liefert

$$0 = \langle v_{k+1}, v_j \rangle = \lambda_{k+1\,k+1} \langle w_{k+1}, v_j \rangle + \mu_{j\,k+1}$$

oder

$$\mu_{j\,k+1} = -\lambda_{k+1\,k+1} \langle w_{k+1}, v_j \rangle. \tag{26}$$

Also sind die $\mu_{j\,k+1}$ und damit die $\lambda_{j\,k+1}$ für $j \leq k$ durch $\lambda_{k+1\,k+1}$ eindeutig bestimmt, d.h. die

$$\lambda_{1\,k+1}, \ldots, \lambda_{k+1\,k+1}$$

sind bis auf einen konstanten Faktor eindeutig bestimmt. Die Forderung $\|v_{k+1}\| = 1$ bestimmt diesen Faktor eindeutig bis aufs Vorzeichen, und die Forderung $\lambda_{k+1\,k+1} > 0$ legt auch noch das Vorzeichen fest.

Nun zeigen wir die Existenz. Inspiriert von (25) und (26) definieren wir zunächst

$$\boxed{\tilde{v}_{k+1} := w_{k+1} - \sum_{j=1}^{k} \langle w_{k+1}, v_j \rangle v_j} \tag{27}$$

und erhalten

$$\langle \tilde{v}_{k+1}, v_j \rangle = 0 \quad \text{für alle } j \in \{1, \ldots, k\}.$$

Weil

$$w_{k+1} \notin \operatorname{Spann}(w_1, \ldots, w_k) = \operatorname{Spann}(v_1, \ldots, v_k),$$

ist $\tilde{v}_{k+1} \neq 0$. Setzen wir

$$\boxed{v_{k+1} = \frac{1}{\|\tilde{v}_{k+1}\|}\, \tilde{v}_{k+1},} \tag{28}$$

so ist das also ein auf $v_1, \ldots, v_k$ orthogonaler Einheitsvektor und $\lambda_{k+1\,k+1} = \frac{1}{\|\tilde{v}_{k+1}\|} > 0$.
□

Die beiden Gleichungen (27) und (28) beschreiben induktiv das Orthonormalisierungsverfahren.

Korollar 5. *Jeder euklidische Vektorraum besitzt eine orthonormale Basis.*

Beachten Sie, daß wegen Lemma 9 die Darstellung eines Vektors bezüglich einer Orthonormalbasis besonders einfach zu finden ist:

$$v = \sum_{i=0}^{n} \langle v, v_i \rangle v_i.$$

Definition 37 (Orthogonalprojektion). Seien $(V, \langle ., . \rangle)$ ein euklidischer Vektorraum und $U \subset V$ ein Untervektorraum. Sei $(u_1, \ldots, u_r)$ eine Orthonormalbasis von U. Dann definiert

$$P_U(v) := \sum_{i=1}^{r} \langle v, u_i \rangle u_i$$

eine lineare Abbildung $P_U : V \to U$, die unabhängig von der gewählten Orthonormalbasis ist. Sie heißt die Orthogonalprojektion auf U.

Beweis der Unabhängigkeit. Sei $(\tilde{u}_1, \ldots, \tilde{u}_r)$ eine weitere Orthonormalbasis von U. Dann gilt für alle $i \in \{1, \ldots, r\}$

$$u_i = \sum_{j=1}^{r} \langle u_i, \tilde{u}_j \rangle \tilde{u}_j, \quad \tilde{u}_j = \sum_{i=1}^{r} \langle \tilde{u}_j, u_i \rangle u_i.$$

Daher ist

$$\begin{aligned}
\sum_{i=1}^{r} \langle v, u_i \rangle u_i &= \sum_{i=1}^{r} \langle v, u_i \rangle \sum_{j=1}^{r} \langle u_i, \tilde{u}_j \rangle \tilde{u}_j \\
&= \sum_{i=1}^{r} \sum_{j=1}^{r} \langle v, u_i \rangle \langle u_i, \tilde{u}_j \rangle \tilde{u}_j = \sum_{i=1}^{r} \sum_{j=1}^{r} \langle v, u_i \langle u_i, \tilde{u}_j \rangle \rangle \tilde{u}_j \\
&= \sum_{j=1}^{r} \langle v, \sum_{i=1}^{r} u_i \langle u_i, \tilde{u}_j \rangle \rangle \tilde{u}_j = \sum_{j=1}^{r} \langle v, \tilde{u}_j \rangle \tilde{u}_j.
\end{aligned}$$

Mit der Definiton von P_U folgt daraus die Unabhängigkeit von der Wahl der ON-Basis in U. $\square$

Satz 25. *Seien $(V, \langle .,. \rangle)$ ein euklidischer Vektorraum und $U \subset V$ ein Untervektorraum. Dann gilt*

$$P_U|U = Id. \tag{29}$$
$$P_U^2 := P_U \circ P_U = P_U. \tag{30}$$
$$\text{Bild}\, P_U = U, \tag{31}$$
$$\text{Kern}\, P_U = \{v \in V \mid \langle v, u \rangle = 0 \textit{ für alle } u \in U\}. \tag{32}$$

Man schreibt $U^\perp := \text{Kern}\, P_U$ *und nennt das* das orthogonale Komplement von U. *Es gilt*

$$P_{U^\perp} = \text{Id} - P_U \tag{33}$$

und

$$(U^\perp)^\perp = U. \tag{34}$$

Weiter ist

$$V = U \oplus U^\perp. \tag{35}$$

Beweis. Sei $(u_1, \dots, u_r)$ eine ON-Basis von U. Zu (29). Für $u \in U$ ist

$$u = \sum_{i=1}^r \langle u, u_i \rangle u_i = P_U(u).$$

Zu (30). Nach (29) ist

$$P_U^2(v) = P_U(\underbrace{P_U(v)}_{\in U}) = P_U(v).$$

Zu (31). Offenbar ist $\text{Bild}\, P_U \subset U$, und aus (29) folgt $U \subset \text{Bild}\, P_U$.

Zu (32). Weil $(u_1, \dots, u_r)$ eine ON-Basis von U ist, ist

$$P_U(v) = \sum_{i=1}^r \langle v, u_i \rangle u_i = 0,$$

äquivalent zu

$$\langle v, u_i \rangle = 0 \quad \text{für alle } i \in \{1, \dots, r\}.$$

Dann ist aber für alle $u \in U$

$$\langle v, u \rangle = \langle v, \sum_{i=1}^r \langle u, u_i \rangle u_i \rangle \sum_{i=1}^r \langle u, u_i \rangle \langle v, u_i \rangle = 0.$$

Also

$$P_U(v) = 0 \implies \langle v, u \rangle = 0 \quad \text{für alle } u \in U.$$

Die Umkehrung ist klar.

Zu (33). Nach der Dimensionsformel und (31), (32) ist

$$\dim U^\perp = \dim \text{Kern}\, P_U = n - \dim U = n - r.$$

Sei $(u_{r+1},\dots,u_n)$ eine ON-Basis von $U^\perp$. Dann ist $(u_1,\dots,u_n)$ eine ON-Basis von V und es gilt für $v\in V$:

$$P_{U^\perp}(v)=\sum_{i=r+1}^n\langle v,u_i\rangle u_i=\sum_{i=1}^n\langle v,u_i\rangle u_i-\sum_{i=1}^r\langle v,u_i\rangle u_i=\mathrm{Id}(v)-P_U(v).$$

Zu (34). Es gilt

$$(U^\perp)^\perp=\mathrm{Kern}\,P_{U^\perp}=\mathrm{Kern}(\mathrm{Id}-P_U)=\{v|v=P_U(v)\}=U.$$

Für die letzte Gleichung beachte, daß $u=P_U(u)$ für alle $u\in U$ nach (31), daß aber andrerseits aus $v=P_U(v)$ natürlich folgt, daß $v\in\mathrm{Bild}\,P_U=U$.

Zu (35). Ist $v\in U\cap U^\perp$, so ist $\langle v,u\rangle=0$ für alle $u\in U$, also auch für $u=v$. Also ist $\langle u,u\rangle=0$ und daher u=0. Es bleibt also zu zeigen, daß $V=U+U^\perp$. Aber es ist

$$v=\underbrace{P_U(v)}_{\in U\text{ nach (31)}}+\underbrace{(Id-P_U)(v)}_{\in U^\perp\text{ nach (33)}}.$$

□

Satz 26 (Dualität). *Ist $(V,\langle .,.\rangle)$ ein euklidischer Vektorraum, so ist*

$$j:V\to V^*,v\mapsto\langle .,v\rangle$$

ein Isomorphismus. Im euklidischen Fall sind also V und V^* kanonisch *isomorph.*

Beweis. Weil $\langle .,.\rangle$ im ersten Argument linear ist, ist $\langle .,v\rangle\in V^*$.

Weil $\langle .,.\rangle$ im zweiten Argument linear ist, ist j linear.

Und weil $j(v)(v)=\langle v,v\rangle>0$ für $v\neq 0$, ist j injektiv, also ein Isomorphismus.

□

Korollar 6. *Seien $(V\langle .,.\rangle)$ ein euklidischer Vektorraum und $w_1,w_2\in V$. Aus*

$$\langle v,w_1\rangle=\langle v,w_2\rangle\quad\textit{für alle } v\in V$$

folgt

$$w_1=w_2.$$

Bemerkung: Es genügt, wenn $\langle v,w_1\rangle=\langle v,w_2\rangle$ für alle v einer Basis von V.

Beweis. Nach Voraussetzung ist $j(w_1)=j(w_2)$. Aber j ist ein Isomorphismus. □

Korollar 7. *Mit den Bezeichnungen des Satzes gilt: Ist $(v_1,\dots,v_n)$ eine Orthonormalbasis von V und $(v_1^*,\dots,v_n^*)$ die duale Basis, so gilt*

$$j(v_i)=v_i^*\textit{ für alle } i\in\{1,\dots,n\}.$$

Beweis.

$$j(v_i)(v_j)=\langle v_j,v_i\rangle=\delta_{ij}=v_i^*(v_j).$$

□

12 Endomorphismen in euklidischen Vektorräumen

In diesem Abschnitt bezeichne $(V, \langle ., . \rangle)$ einen n-dimensionalen euklidischen Vektorraum.

Definition 38. Ist $f \in \mathrm{End}(V)$ und $f^* : V^* \to V^*$ die duale Abbildung, so heißt

$$f^\# := j^{-1} \circ f^* \circ j : V \to V$$

die Adjungierte von f. Sie ist charakterisiert durch

$$\langle f(v), w \rangle = \langle v, f^\#(w) \rangle \quad \text{für alle } v, w \in V. \tag{36}$$

Beweis.

$$\begin{aligned} \langle v, f^\#(w) \rangle &= j(f^\#(w))(v) = j(j^{-1} \circ f^* \circ j(w))(v) \\ &= f^*(j(w))(v) = j(w)(f(v)) = \langle f(v), w \rangle. \end{aligned}$$

□

Wenn man vermöge $j : V \to V^*$ die beiden Räume identifiziert, so entsprechen sich die adjungierte und die duale Abbildung.

$$\begin{array}{ccccc} & V^* & \underset{f^*}{\rightarrow} & V^* & \\ j & \uparrow\cong & & \cong\uparrow & j \\ & & & & \\ & V & \underset{f^\#}{\rightarrow} & V & \end{array}$$

Es ist deshalb üblich, auch die Adjungierte von $f \in \mathrm{End}(V)$ mit f^* statt $f^\#$ zu bezeichnen, der Definitionsbereich entscheidet dann, um welche der beiden Abbildungen es sich handelt. Wir folgen im weiteren dieser Konvention:

$$\boxed{f^* : V \to V \quad \text{bezeichne die Adjungierte von } f.}$$

Satz 27. *Seien* $\mathbf{v} = (v_1, \ldots, v_n)$ *eine Orthonormalbasis von* V *und* $f \in \mathrm{End}(V)$. *Dann gilt für die Darstellungsmatrix*

$$A_{\mathbf{v}}^{\mathbf{v}}(f) = (\langle v_i, f(v_j) \rangle)_{i,j=1,\ldots,n}.$$

Beweis. Sei (a_{ij}) die Darstellungsmatrix, also

$$f(v_j) = \sum_{i=1}^{n} a_{kj} v_k.$$

Durch Skalarmultiplikation mit v_i folgt die Behauptung. □

Korollar 8. *Bezüglich einer Orthonormalbasis* $\mathbf{v}$ *haben* $f \in \mathrm{End}(V)$ *und die Adjungierte* f^* *transponierte Darstellungsmatrizen:*

$$\boxed{A_{\mathbf{v}}^{\mathbf{v}}(f^*) = \left(A_{\mathbf{v}}^{\mathbf{v}}(f)\right)^T.}$$

Beweis.
$$\langle v_i, f^*(v_j)\rangle = \langle f(v_i), v_j\rangle = \langle v_j, f(v_i)\rangle.$$
□

Definition 39 (Orthogonale Abbildungen). Ein Endomorphismus $f : V \to V$ heißt *orthogonal* oder *eine orthogonale Abbildung*, falls

$$\langle f(v), f(w)\rangle = \langle v, w\rangle \qquad \text{für alle } v, w \in V.$$

Die Menge der orthogonalen Abbildungen von V wird mit $\mathrm{O}(V)$ bezeichnet.

Bemerkung. Orthogonale Endomorphismen erhalten das Skalarprodukt und damit auch die daraus abgeleiteten Größen: Die Länge von Vektoren und der Winkel zwischen Vektoren bleiben bei orthogonalen Abbildungen erhalten. Insbesondere sind sie injektiv, also

$$\mathrm{O}(V) \subset \mathrm{Aut}(V).$$

Satz 28 (Längentreue Endomorphismen). *Sei* $f \in \mathrm{End}(V)$ *mit*

$$\|f(v)\| = \|v\| \quad \textit{für alle } v \in V.$$

Dann ist $f \in \mathrm{O}(V)$.

Beweis. Aus $\|f(v+w)\|^2 = \|v+w\|^2$ folgt

$$\|f(v)\|^2 + \|f(w)\|^2 + 2\langle f(v), f(w)\rangle = \|v\|^2 + \|w\|^2 + 2\langle v, w\rangle,$$

also
$$\langle f(v), f(w)\rangle = \langle v, w\rangle.$$
□

Bemerkung. Ist allgemeiner $f : V \to V$ eine beliebige längentreue Abbildung, also

$$\|f(v) - f(w)\| = \|v - w\| \text{ für alle } v, w \in V,$$

und gilt überdies

$$f(0) = 0,$$

so ist f linear. Nach dem vorangehenden Satz ist es also ein orthogonaler Endomorphismus.

Beweis. Zunächst folgt mit $w = 0$, daß

$$\|f(v)\| = \|v\|$$

für alle v. Weiter ist

$$\begin{array}{c} \|f(v) - f(w)\|^2 = \|f(v)\|^2 + \|f(w)\|^2 - 2\langle f(v), f(w)\rangle \\ \| \\ \|v - w\|^2 = \|v\|^2 + \|w\|^2 - 2\langle v, w\rangle, \end{array}$$

also
$$\langle f(v), f(w)\rangle = \langle v, w\rangle$$

für alle v, w. Dann ist aber

$$\begin{aligned}
\|f(\lambda v + \mu w) - \lambda f(v) - \mu f(w)\|^2 &= \|f(\lambda v + \mu w)\| + \lambda^2 \|f(v)\|^2 + \mu^2 \|f(w)\| \\
&\quad - 2\lambda \langle f(\lambda v + \mu, f(v)\rangle - 2\mu \langle f(\lambda v + \mu w), f(w)\rangle \\
&\quad + 2\langle f(v), f(w)\rangle \\
&= \|\lambda v + \mu w\| + \lambda^2 \|v\|^2 + \mu^2 \|w\| \\
&\quad - 2\lambda \langle \lambda v + \mu w, v\rangle - 2\mu \langle \lambda v + \mu w, w\rangle \\
&\quad + 2\langle f(v), f(w)\rangle \\
&= \|(\lambda v + \mu w) - \lambda v - \mu w\|^2 \\
&= 0.
\end{aligned}$$

□

Satz 29 (Die orthogonale Gruppe). *Die Menge* $\mathrm{O}(V)$ *ist bezüglich der Komposition eine Gruppe, die orthogonale Gruppe von* $(V, \langle ., . \rangle)$ *mit* Id *als neutralem Element. Man nennt* $\mathrm{O}(V)$ *die orthogonale Gruppe von* $(V, \langle ., . \rangle)$.

Beweis. Wir müssen zeigen, daß die Komposition von orthogonalen Abbildungen und die Inverse einer orthogonalen Abbildung wieder orthogonal sind. Beides ist trivial. Das Assoziativgesetz gilt allgemein für die Komposition von Abbildungen. □

Satz 30. *Sei* f^* *die Adjungierte von* $f \in \mathrm{End}(V)$. *Dann gilt*

$$f \in \mathrm{O}(V) \iff f f^* = f^* f = \mathrm{Id}\,.$$

Beweis.

$$\begin{aligned}
f \in \mathrm{O}(V) &\iff \forall_{v,w} \langle f(v), f(w)\rangle = \langle v, w\rangle \\
&\iff \forall_{v,w} \langle v, f^*(f(w))\rangle = \langle v, w\rangle \\
&\iff \forall_{v,w} \langle v, f^*(f(w)) - w\rangle = 0
\end{aligned}$$

Einsetzen von $v = f^*(f(w)) - w$ zeigt, daß dies genau dann der Fall ist, wenn

$$f^*(f(w)) - w = 0 \quad \text{für alle } w \in V,$$

d.h. wenn $f^* = f^{-1}$. □

Korollar 9. *Seien* $f \in \mathrm{End}(V)$ *und* $A = (a_{ij}) = A^{\mathbf{v}}_{\mathbf{v}}(f)$ *seine Darstellungsmatrix bezüglich einer Orthonormalbasis. Dann sind äquivalent:*

1. $f \in \mathrm{O}(V)$
2. *Für alle* $i, k \in \{1, \ldots, n\}$ *ist*

$$\sum_{j=1}^{n} a_{ij} a_{kj} = \delta_{ik},$$

d.h. die Zeilen von A *sind orthonormal im* $\mathbb{R}^n$ *bezüglich des kanonischen Skalarproduktes.*

3. Für alle $i, k \in \{1, \dots, n\}$ ist

$$\sum_{j=1}^{n} a_{ji} a_{jk} = \delta_{ik},$$

d.h. die Spalten von A sind orthonormal im $\mathbb{R}^n$ bezüglich des kanonischen Skalarproduktes.

Beweis. Es gilt

$$A^T A = \left(\sum_{j=1}^{n} a_{ij} a_{kj} \right), \qquad AA^T = \left(\sum_{j=1}^{n} a_{ji} a_{jk} \right).$$

Weil aber A^T die Darstellungsmatrix von f^* ist, folgt daraus die Behauptung. □

Definition 40. Eine reelle (n, n)-Matrix A heißt *orthogonal*, wenn ihre Zeilen (oder Spalten) orthonormal sind, d.h. wenn

$$A^T A = AA^T = E. \qquad E = (\delta_{ij}) = \text{Einheitsmatrix}$$

Die Menge der orthogonalen (n, n)-Matrizen bildet eine Gruppe, die *orthogonale Gruppe* $\mathrm{O}(n)$ *der Ordnung* n.

(Der Begriff Ordnung weicht in diesem Zusammenhang ab von dem, was man in der Gruppentheorie darunter versteht: n ist nicht etwa die Anzahl der Elemente von $\mathrm{O}(n)$.)

Beispiel 55 (Orthogonale Abbildungen im $\mathbb{R}^2$). Für $\phi \in \mathbb{R}$ sind

$$A_\pm := \begin{pmatrix} \cos\phi & \mp\sin\phi \\ \sin\phi & \pm\cos\phi \end{pmatrix} \in \mathrm{O}(2).$$

Als lineare Abbildung des $\mathbb{R}^2$ in sich beschreibt A_+ gerade die Drehung um den Ursprung um den Winkel ϕ, während A_- eine Spiegelung an der x-Achse mit nachfolgender Drehung um ϕ ist:

$$A_- = \begin{pmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} = A_+ \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

Ist umgekehrt $A \in \mathrm{O}(2)$, so ist der erste Spaltenvektor ein Einheitsvektor. Nach Analysis läßt der sich schreiben als $\begin{pmatrix} \cos\phi \\ \sin\phi \end{pmatrix}$ und ein dazu senkrechter Einheitsvektor ist eindeutig bis aufs Vorzeichen

$$\begin{pmatrix} \mp\sin\phi \\ \pm\cos\phi \end{pmatrix}.$$

Also ist jedes $A \in \mathrm{O}(2)$ von obiger Form.

□

Beispiel 56 (Orthogonale Abbildungen des $\mathbb{R}^3$). Sei $f \in \mathrm{O}(\mathbb{R}^3)$. Wir werden später zeigen, daß es dann einen Einheitsvektor $v_1 \in \mathbb{R}^3$ gibt, so daß

$$f(v_1) = \pm v_1$$

ist. Wir ergänzen v_1 zu einer ON-Basis $\mathbf{v}(v_1, v_2, v_3)$ des $\mathbb{R}^3$. Dann hat die zugehörige Darstellungsmatrix von f die Gestalt

$$A_{\mathbf{v}}^{\mathbf{v}}(f) = \begin{pmatrix} \pm 1 & 0 \; 0 \\ 0 & \\ 0 & B \end{pmatrix}$$

Dabei muß die $(2,2)$-Matrix B wieder orthogonal sein, also von der Form im letzten Beispiel. Daher ist eine orthogonale Abbildung des $\mathbb{R}^3$ eine Drehung um eine Achse $\mathbb{R}v_1$, eventuell noch komponiert mit Spiegelungen:

$$A_{\mathbf{v}}^{\mathbf{v}}(f) = \begin{pmatrix} \pm 1 & 0 & 0 \\ 0 & \cos\phi & \mp\sin\phi \\ 0 & \sin\phi & \pm\cos\phi \end{pmatrix} = \begin{pmatrix} \pm 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\phi & -\sin\phi \\ 0 & \sin\phi & \cos\phi \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \pm 1 \end{pmatrix}.$$

□

Definition 41. Seien $(V\langle .,.\rangle)$ ein euklidischer Vektorraum und $f \in \mathrm{End}(V)$.

1. f heißt *selbstadjungiert* oder *symmetrisch*, falls
$$f^* = f.$$
2. f heißt *schiefadjungiert* oder *schiefsymmetrisch*, falls
$$f^* = -f.$$
3. f heißt *normal*, falls
$$f^*f = ff^*$$

Satz 31. *Seien $(V\langle .,.\rangle)$ ein euklidischer Vektorraum und $f \in \mathrm{End}(V)$ und A die Darstellungsmatrix von f bezüglich einer Orthonormalbasis von V. Dann gilt*

$$\begin{aligned} f \textit{selbstadjungiert} &\iff A^T = A \iff \forall_{v,w\in V}\langle f(v), w\rangle = \langle v, f(w)\rangle, \\ f \textit{schiefadjungiert} &\iff A^T = -A \iff \forall_{v,w\in V}\langle f(v), w\rangle = -\langle v, f(w)\rangle, \\ f \textit{normal} &\iff A^TA = AA^T. \end{aligned}$$

Beweis. Triviale Folge der Definiton und des Korollars 8. □

Beispiel 57 (Orthogonalprojektionen). Ist P_U die Orthogonalprojektion auf einen Unterraum $U \subset V$, so ist P_U selbstadjungiert, denn

$$\langle P_U(v), w\rangle = \langle \sum_{i=1}^k \langle v, u_i\rangle u_i, w\rangle = \sum_{i=1}^k \langle v, u_i\rangle\langle u_i, w\rangle$$

ist offensichtlich symmetrisch in v und w. Umgekehrt sind Orthogonalprojektionen durch die Selbstadjungiertheit in Verbindung mit $f^2 = f$ charakterisiert:

□

Satz 32 (Orthogonalprojektionen). *Sei $f \in \mathrm{End}(V)$ mit*

$$f^2 = f, \qquad f^* = f.$$

Dann ist f eine Orthogonalprojektion, nämlich

$$f = P_{\mathrm{Bild}\, f}.$$

Beweis. Zunächst ist
$$\operatorname{Kern} f \perp \operatorname{Bild} f.$$
Ist nämlich $v \in \operatorname{Kern} f$, also $f(v) = 0$, und $u \in \operatorname{Bild} f$, also $u = f(w)$ für irgend ein $w \in V$, so folgt
$$\langle v, u\rangle = \langle v, f(w)\rangle = \langle f^*(v), w\rangle = \langle f(v), w\rangle = \langle 0, w\rangle = 0.$$
Nach dem Dimensionssatz ist daher
$$\operatorname{Kern} f = (\operatorname{Bild} f)^\perp. \tag{37}$$
Weiter gilt für $u = f(w) \in \operatorname{Bild} f$
$$f(u) = f(f(w)) = f^2(w) = f(w) = u,$$
also $f|_U = \mathrm{Id}$. Sei nun $u_1, \ldots, u_k$ eine ON-Basis von $\operatorname{Bild} f$, die wir zu einer ON-Basis $u_1, \ldots, u_n$ von V ergänzen. Nach (37) ist dann $u_{k+1}, \ldots, u_n$ eine ON-Basis von $\operatorname{Kern} f$. Wir finden für $v \in V$:
$$\begin{aligned} f(v) &= f(\sum_{i=1}^n \langle v, u_i\rangle u_i) = \sum_{i=1}^n \langle v, u_i\rangle f(u_i) \\ &= \sum_{i=1}^k \langle v, u_i\rangle f(u_i) = \sum_{i=1}^k \langle v, u_i\rangle g u_i \\ &= P_{\operatorname{Bild} f}(v). \end{aligned}$$
□

Zur Bedeutung von selbstadjungierten, schiefadjungierten und normalen Endomorphismen.

Zunächst sind selbstadjungierte, schiefadjungierte und orthonormale Endomorphismen sämtlich normal, und „Normalität" ist im wesentlichen ein technischer Begriff, der es erlaubt, insbesondere allen diesen Klassen gemeinsame Eigenschaften zu behandeln. Wir kommen bei der Normalformtheorie darauf zurück.

Schief- und Selbstadjungiertheit sind aus mehreren Gründen wichtige Phänomene. Ich gebe für jede der beiden Eigenschaften *eine* wichtige Bedeutung aus physikalischen Modellen.

Wenn wir eine Drehung eines euklidischen Raums V um 0 als einen zeitabhängigen Prozeß beschreiben wollen, können wir eine Abbildung $\mathbb{R} \to \mathrm{O}(V), t \mapsto f_t$ betrachten. Wir nehmen an daß $f_0 = \mathrm{Id}$, d.h. zur Zeit t ist alles an seinem Platz, und zur Zeit t ist der Vektor v in die Position $f_t(v)$ gedreht. Die Geschwindigkeit dieser Bewegung ist gegeben durch
$$\frac{d}{dt} f_t(v) =: \dot{f}_t v$$
Aus $\langle f_t(v), f_t w\rangle = \langle v, w\rangle$ für alle t folgt aber durch Differentiation nach der Produktregel (vgl. Analysis)
$$\langle \dot{f}_0(v), f_0(w)\rangle + \langle f_0(v), \dot{f}_0(w)\rangle = 0,$$
also
$$< \dot{f}_0(v), w > + \langle v, \dot{f}_0(w)\rangle = 0.$$
Das heißt $\dot{f}_0 \in \operatorname{End}(V)$ ist schiefadjungiert. Physikalisch gesprochen bedeutet das:

Infinitesimale Drehungen werden durch schiefadjungierte Endomorphismen beschrieben.

Die Selbstadjungiertheit von $f \in \operatorname{End}(V)$ bedeutet, daß durch
$$\alpha(v, w) := \langle f(v), w\rangle$$
eine *symmetrische* Bilinearform definiert wird. Solche Bilinearformen spielen zum Beispiel als Trägheitstensoren oder Spannungstensoren in der Mechanik eine wichtig Rolle.

13 Unitäre Vektorräume

In diesem Abschnitt betrachten wir Vektorräume über dem Körper $\mathbb{C}$ der komplexen Zahlen. Wir bezeichnen mit einem Querstrich die komplexe Konjugation: Für $x, y \in \mathbb{R}$ ist

$$\overline{x+iy} = x - iy.$$

Definition 42 (Hermitesche Formen). Sei V ein $\mathbb{C}$-Vektorraum.

1. Eine Abbildung
$$\langle .,. \rangle : V \times V \to \mathbb{C}, (v, w) \mapsto \langle v, w \rangle$$
heißt Sesquilinearform (sesquilinear=$1\frac{1}{2}$-linear), wenn gilt:
Für alle $v, w, v_1, v_2, w_1, w_2 \in V$ und $\lambda \in \mathbb{C}$ ist
$$\langle v_1 + v_2, w \rangle = \langle v_1, w \rangle + \langle v_2, w \rangle$$
$$\langle v, w_1 + w_2 \rangle = \langle v, w_1 \rangle + \langle v, w_2 \rangle$$
$$\langle \lambda v, w \rangle = \lambda \langle v, w \rangle, \quad \langle v, \lambda w \rangle = \bar{\lambda} \langle v, w \rangle.$$

2. Eine Sesquilinearform $\langle .,. \rangle$ auf V heißt *eine hermitesche Form*, wenn für alle $v, w \in V$
$$\langle v, w \rangle = \overline{\langle w, v \rangle}.$$

3. Eine hermitesche Form $\langle .,. \rangle$ auf V heißt ein *hermitesches Skalarprodukt*, wenn sie überdies *positiv definit* ist, d.h. wenn für alle $v \in V \setminus \{0\}$
$$\langle v, v \rangle > 0.$$
Beachten Sie, daß $\langle v, v \rangle$ nach Punkt 2 reell ist.

4. Ein Paar $(V, \langle .,. \rangle)$ aus einem $\mathbb{C}$-Vektorraum V mit einem hermiteschen Skalarprodukt $\langle .,. \rangle$ heißt ein *unitärer (Vektor)raum*.

5. Ist $\langle .,. \rangle$ ein hermitesches Skalarprodukt auf V, so definieren wir die dadurch induzierte *Norm* durch
$$\|v\| := \sqrt{\langle v, v \rangle} \text{ für alle } v \in V.$$
Wir nennen $\|v\|$ auch die *Länge* des Vektors $\|v\|$.

Beispiel 58. Im $\mathbb{C}^n$ definiert
$$\langle v, w \rangle := \sum_{k=1}^{n} v_i \overline{w_i}$$
das kanonische hermitesche Skalarprodukt.

□

Bemerkungen. Wie im euklidischen Fall beweist oder definiert man

- die Cauchy-Schwarzsche Ungleichung,
- die Dreiecksungleichung,
- die Begriffe *orthogonal, orthonormal*,
- das E. Schmidtsche Orthonormalisierungsverfahren und die Existenz von Orthonormalbasen in endlich-dimensionalen unitären Räumen,

- die Existenz und Eigenschaften von Orthogonalprojektionen.
  (Beachte: $P_U(v) = \sum_{i=1}^k \langle v, u_i\rangle u_i$ ist in v linear!)

Ein hermitesches Skalarprodukt ist durch die Norm eindeutig bestimmt: Wie im euklidischen Fall zeigt man

$$\|v+w\|^2 = \|v\|^2 + \|w\|^2 + \langle v, w\rangle + \langle w, v\rangle = \|v\|^2 + \|w\|^2 + 2\,\mathrm{Re}\langle v, w\rangle.$$

Aber das reicht aus. Wenn man für alle v, w den Realteil von $\langle v, w\rangle$ kennt, kennt man auch den Imaginärteil:

$$\mathrm{Im}\langle v, w\rangle = -\,\mathrm{Re}(i\langle v, w\rangle) = \mathrm{Re}\langle v, iw\rangle.$$

Darstellungsmatrix. Aufpassen muß man bei der Darstellungsmatrix bezüglich einer orthonormalen Basis $(v_1, \ldots, v_n)$: Aus

$$f(v_j) = \sum_{i=1}^n a_{ij} v_i \tag{38}$$

folgt durch *Rechts*multiplikation mit v_i

$$\boxed{\langle f(v_j), v_i\rangle = a_{ij},}$$

und das ist $\overline{\langle v_i, f(v_j)\rangle}$. Am besten merkt man sich (38) und leitet sich die Formel für die a_{ij} daraus her.

Dualität und Adjungierte. Ein fundamentaler Unterschied zum euklidischen Fall ist die Tatsache, daß

$$j : V \to V^*,\ w \mapsto \langle ., w\rangle$$

zwar eine Bijektion, aber nicht linear ist:

$$j(\lambda w) = \bar{\lambda} j(w).$$

j ist ein sogenannter *antilinearer* Isomorphismus.

Die Adjungierte f^* eines Endomorphismus ist auch im unitären Fall charakterisiert durch

$$\langle f(v), w\rangle = \langle v, f^*(w)\rangle.$$

Sie entspricht der dualen Abbildung unter $j : V \to V^*$ (Beachten Sie, das $j^{-1} \circ f^* \circ j$ wieder $\mathbb{C}$-linear ist.) Daher ist die Bezeichnung mit demselben Symbol f^* einigermaßen entschuldbar. Für die Darstellungsmatrix $A^* = (a^*_{ij})$ von f^* bezüglich einer Orthonormalbasis gilt

$$\boxed{a^*_{ij} = \overline{a_{ji}},}$$

also

$$\boxed{A^* = \bar{A}^T.}$$

Von euklidischen Fall unterscheidet sich das durch die Konjugation.

Beweis.

$$a^*_{ij} = \langle f^*(v_j), v_i\rangle = \langle v_j, f(v_i)\rangle = \overline{\langle f(v_i), v_j\rangle} = \overline{a_{ji}}.$$

□

Selbstadjungierte und schiefadjungierte Endomorphismen werden definiert wie im euklidischen Fall.

Unitäre Endomorphismen. Wir kommen nun zum Analogon orthogonaler Endomorphismen im unitären Fall.

Definition 43. Seien $(V, \langle ., . \rangle)$ ein unitärer Vektorraum und $f \in \operatorname{End}(V)$. f heißt *unitär*, falls

$$\langle f(v), f(w) \rangle = \langle v, w \rangle \quad \text{für alle } v, w \in V.$$

Wie im euklidischen beweist man:

Satz 33. *Seien $(V, \langle ., . \rangle)$ ein n-dimensionaler unitärer Vektorraum und $f \in \operatorname{End}(V)$. Sei A die Darstellungsmatrix von f bezüglich einer Orthonormalbasis. Wir bezeichnen mit f^* die Adjungierte und setzen $A^* = \bar{A}^T$. Dann sind folgende Aussagen äquivalent:*

1. *f ist unitär*
2. *$\|f(v)\| = \|v\|$ für alle $v \in V$.*
3. *$ff^* = f^*f = \mathrm{Id}$.*
4. *$AA^* = A^*A = E$.*
5. *Die Zeile von A bilden ein Orthonormalsystem im $\mathbb{C}^n$.*
6. *Die Spalten von A bilden ein Orthonormalsystem im $\mathbb{C}^n$.*

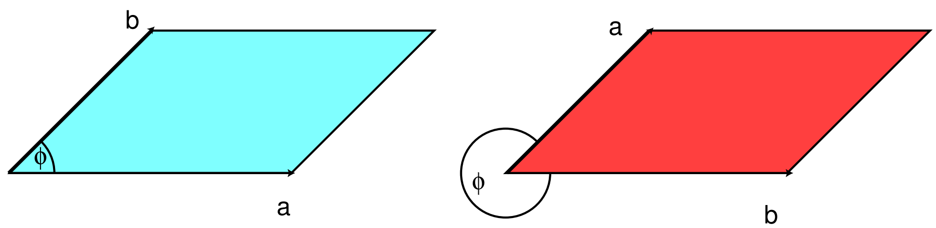
Satz 34 (Die unitäre Gruppe). *Die Menge $U(V)$ der unitären Endomorphismen eines unitären Vektorraumes $(V, \langle ., . \rangle)$ bildet eine Gruppe, die* unitäre Gruppe von V.
Die Menge $U(n) = \{A \,|\, A \text{ ist } (n,n)\text{-Matrix mit } A^*A = A^*A = E\}$ bildet eine Gruppe die unitäre Gruppe *der Ordnung n.*

14 Die Determinante

Wir beginnen mit einer geometrischen Motivation und setzen dafür voraus, daß wir die Flächenberechnung von Parallelelogrammen in der Ebene kennen: Die Fläche ist das Produkt aus der Grundseite und der Höhe. Geben wir zwei benachbarte Seiten des Parallelogramms durch Vektoren $a, b \in \mathbb{R}^2$, so ist die Fläche

$$F(a, b) = \|a\| \, \|b\| \sin \phi.$$

Dabei bezeichnet $\|a\|$ die *Länge* des Vektors oder der Seite a, entsprechend $\|b\|$, und ϕ den Winkel zwischen den beiden Vektoren. Ist dieser Winkel $> \pi$, so ist der Flächeninhalt negativ, und so wollen wir ihn auch lassen: *Flächeninhalt* soll *signierten Flächeninhalt*, also Flächeninhalt mit Vorzeichen bedeuten. Das wird deutlich an diesem Bild:



Wir erhalten damit

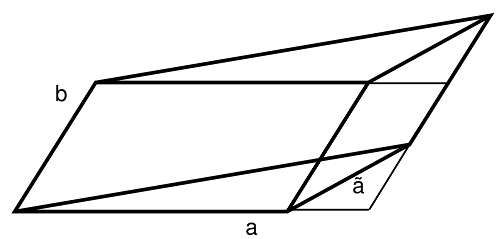
$$F(a, b) = -F(b, a).$$

Offenbar ist für $\lambda \in \mathbb{R}$

$$F(\lambda a, b) = \lambda F(a, b).$$

Machen Sie sich anhand der nebenstehenden Figur klar, daß weiter

$$F(a + \tilde{a}, b) = F(a, b) + F(\tilde{a}, b).$$



Ähnliche Eigenschaften hat das Volumen $V(a, b, c)$ des von drei Vektoren a, b, c im $\mathbb{R}^3$ aufgespannte *Parallelotop*. Es gilt zum Beispiel:

$$\begin{aligned} V(a, b, c) &= V(c, a, b) = -V(b, a, c) \\ V(a + \tilde{a}, b, c) &= V(a, b, c) + V(\tilde{a}, b, c) \\ V(\lambda a, b, c) &= \lambda V(a, b, c) \end{aligned}$$

Wir wollen nun solche „Volumenfunktionen“ verallgemeinern.

Definition 44 (Multilinearformen). Seien V ein n-dimensionaler K-Vektorraum und $k \in \mathbb{N} \backslash \{0\}$. Eine *$k$-Linearform*, kurz eine *k-Form*, auf V ist eine Abbildung

$$\omega : V^k = \underbrace{V \times \ldots \times V}_{k \text{ Faktoren}} \to K$$

mit folgenden Eigenschaften:

Für alle $i \in \{1, \ldots, k\}$, $v_1, \ldots, v_k, \tilde{v}_i \in V$ und $\lambda \in K$ gelten

$$\begin{aligned} \omega(v_1, \ldots, v_i + \tilde{v}_i, \ldots, v_k) &= \omega(v_1, \ldots, v_i, \ldots, v_k) + \omega(v_1, \ldots, \tilde{v}_i, \ldots, v_k), \\ \omega(v_1, \ldots, \lambda v_i, \ldots, v_k) &= \lambda \omega(v_1, \ldots, v_i, \ldots, v_k). \end{aligned}$$

Die Abbildung ω ist also in jedem Argument linear. Wir kennen schon 2-Linearformen, nämlich *Bilinearformen*, und 1-Linearformen, nämlich *Linearformen*. Wenn man k nicht spezifizieren will, spricht man von *Multilinearformen*.

Eine k-Linearform ω heißt *symmetrisch*, wenn ihr Wert sich bei der Vertauschung zweier beliebiger Argumente nicht ändert. Sie heißt *alternierend* oder *schiefsymmetrisch*, falls

$$\omega(v_1, \dots, v_k) = 0,$$

wenn immer zwei Argumente gleich sind: $v_i = v_j$ für ein Paar $i \neq j$. Zur Rechtfertigung dieser etwas überraschenden Definition beweisen wir

Lemma 10. *Der Wert einer alternierenden k-Linearform ω ändert sein Vorzeichen, wenn man zwei Argumente vertauscht.*

Daher kommt der Name *alternierend*. In Körpern der Charakteristik $\neq 2$ gilt auch die Umkehrung: Ändert ω bei Vertauschung zweier Argumente sein Vorzeichen, so ist es im Sinne der obigen Definition alternierend.

Beweis des Lemmas. Seien $i < j$ und $v_1, \dots, v_k \in V$. Dann ist

$$\begin{aligned}\omega(v_1, \dots, &v_i + v_j, \dots, v_j + v_i, \dots, v_k) \\ &= \omega(v_1, \dots, v_i, \dots, v_i, \dots, v_k) + \omega(v_1, \dots, v_i, \dots, v_j, \dots, v_k) \\ &\quad + \omega(v_1, \dots, v_j, \dots, v_i, \dots, v_k) + \omega(v_1, \dots, v_j, \dots, v_j, \dots, v_k).\end{aligned}$$

Alle Terme mit zwei gleichen Argumenten sind nach Voraussetzung $= 0$. Also folgt

$$0 = \omega(v_1, \dots, v_i, \dots, v_j, \dots, v_k) + \omega(v_1, \dots, v_j, \dots, v_i, \dots, v_k).$$

□

Beispiel 59. Sei $(V, \langle ., . \rangle)$ ein euklidischer Vektorraum und $f \in \mathrm{End}(V)$. Dann definiert

$$\omega(v, w) := \langle f(v), w \rangle$$

eine Bilinearform, die symmetrisch ist, wenn f selbstadjungiert ist, und die alternierend ist, wenn f schiefadjungiert ist. Zum Beispiel erhält man für den $\mathbb{R}^{2n}$ und $f = f_A$ mit

$$A = \begin{pmatrix} & & & -1 & \dots & 0 \\ & 0 & & \vdots & \ddots & \vdots \\ & & & 0 & \dots & -1 \\ 1 & \dots & 0 & & & \\ \vdots & \ddots & \vdots & & 0 & \\ 0 & \dots & 1 & & & \end{pmatrix}$$

die alternierende Bilinearform

$$\omega(x, y) = \sum_{i=1}^{n} x_i y_{n+i} - x_{n+i} y_i,$$

die in der Hamiltonschen Mechanik eine wichtige Rolle spielt.

□

Lemma 11. *Ist ω eine alternierende k-Form auf V und sind $v_1, \dots, v_k \in V$ linear abhängig, so ist*

$$\omega(v_1, \dots, v_k) = 0.$$

Beweis. Eines der v_i ist eine Linearkombination der anderen. Sei der Einfachheit halber $v_1 = \sum_{j=2}^n \lambda_j v_j$. Dann ist

$$\omega(v_1, \dots, v_k) = \omega(\sum_{j=2}^n \lambda_j v_j, v_2, \dots, v_k) = \sum_{j=2}^n \lambda_j \underbrace{\omega(v_j, v_2, \dots, v_k)}_{=0}$$

□

Satz 35. *Seien V ein n-dimensionaler K-Vektorraum mit Basis $(b_1, \dots, b_n)$ und ω eine alternierende k-Form auf V. Dann gilt*

1. *Ist $k > n$, so ist $\omega = 0$.*
2. *Ist $k = n$, so ist ω durch den Wert*

$$\omega(b_1, \dots, b_n)$$

eindeutig bestimmt.

Beweis. Die erste Behauptung folgt aus den Lemma. Wir zeigen sie hier aber noch einmal. Sei

$$v_j = \sum_{i_j=1}^n x_{i_j j} b_{i_j}.$$

Beachten Sie, daß man den Summationsindex natürlich auch einfach i statt i_j nennen könnte. Aber jetzt brauchen wir alle $v's$ gleichzeitig: Es ist nämlich

$$\omega(v_1, \dots, v_k) = \omega(\sum_{i_1} x_{i_1 1} b_{i_1}, \dots, \sum_{i_k} x_{i_k k} b_{i_k}) = \sum_{i_1} \dots \sum_{i_k} x_{i_1 1} \dots x_{i_k k} \omega(b_{i_1}, \dots, b_{i_k}).$$

Hier verschwinden alle Terme, bei denen in $(b_{i_1}, \dots, b_{i_k})$ zwei der $b's$ gleich sind. Ist $k > n$, so ist das stets der Fall und $\omega(v_1, \dots, v_k) = 0$ für alle $(v_1, \dots, v_k)$. Damit ist 1. bewiesen.

Ist $k = n$, so bleiben genau die Terme, bei denen $(b_{i_1}, \dots, b_{i_n})$ eine Umordnung von $(b_1, \dots, b_n)$ ist. Also gilt

$$\omega(b_{i_1}, \dots, b_{i_n}) = \epsilon_{i_1 \dots i_n} \omega(b_1, \dots, b_n),$$

wobei $\epsilon_{i_1 \dots i_n} = \pm 1$ abhängig davon, wieviele Vertauschungen man braucht, um $(b_{i_1}, \dots, b_{i_n})$ in die richtige Reihenfolge zu bringen. Daraus folgt 2. □

Korollar 10. *Zwei alternierende n-Formen auf einem n-dimensionalen Vektorraum unterscheiden sich nur um einen konstanten Faktor.*

Definition 45 (Determinantenform). Eine *Determinantenform* auf einem n-dimensionalen Vektorraum ist eine nicht-triviale alternierende n-Linearform auf V.

Das vorstehende Korollar zeigt, daß es bis auf konstante Faktoren höchstens eine solche Determinantenform auf V geben kann. Wir zeigen jetzt die Existenz.

Satz 36. *Ist V ein n-dimensionaler K-Vektorraum, $n > 0$, so existiert auf V eine Determinantenform.*

Beweis. Durch vollständige Induktion über n.

$\underline{n = 1}$. Dann ist $\dim V^* = 1$ und jedes $\omega \in V^* \backslash \{0\}$ ist eine Determinantenform.

$\underline{n \to n+1}$. Sei also der Satz bereits für alle n-dimensionalen K-Vektorräume bewiesen. Sei $\tilde{V}$ ein $(n+1)$-dimensionaler K-Vektorraum. Dann gibt es in $\tilde{V}$ zwei Untervektorräume V und W, so daß $\dim V = n$ und

$$\tilde{V} = V \bigoplus W.$$

Man kann zum Beispiel eine Basis $(b_0, \ldots, b_n)$ von $\tilde{V}$ nehmen und definieren

$$W = \operatorname{Spann}(b_0), \qquad V = \operatorname{Spann}(b_1, \ldots, b_n).$$

Dann ist $\dim W = 1$. Nach Induktionsvoraussetzung gibt es Determinantenformen ϕ und ψ auf V und W. Jeder Vektor $\tilde{v} \in \tilde{V}$ läßt sich eindeutig schreiben als

$$\tilde{v} = v + w \text{ mit } v \in V \text{ und } w \in W,$$

und mit dieser Notation definieren wir nun eine $(n+1)$-Form ω auf $\tilde{V}$ durch

$$\omega(v_0 + w_0, \ldots, v_n + w_n) := \sum_{i=0}^{n} (-1)^i \phi(v_0, \ldots, \widehat{v_i}, \ldots, v_n) \psi(w_i). \tag{39}$$

Die Linearität im j-ten Argument folgt so:

Ersetzt man $v_j + w_j$ durch $v_j + w_j + v_j^* + w_j^* = (v_j + v_j^*) + (w_j + w_j^*)$, so sind die Summanden rechts (bis aufs Vorzeichen) von der Form

$$\phi(v_0, \ldots, \widehat{v_j}, \ldots, v_n)\psi(w_j + w_j^*) = \phi(v_0, \ldots, \widehat{v_j}, \ldots, v_n)\psi(w_j) + \phi(v_0, \ldots, \widehat{v_j}, \ldots, v_n)\psi(w_j^*)$$

oder, für $i \neq j$, von der Form

$$\begin{aligned} \phi(v_0, \ldots, \widehat{v_i}, \ldots, v_j + v_j^*, \ldots, v_n)\psi(w_i) = \phi(v_0, \ldots, \widehat{v_i}, \ldots, v_j, \ldots, v_n)\psi(w_i) \\ + \phi(v_0, \ldots, \widehat{v_i}, \ldots, v_j^*, \ldots, v_n)\psi(w_i). \end{aligned}$$

Man kann die Summe rechts in (39) also in zwei Summen auseinanderziehen und findet die Additivität im j-ten Argument. Die Homogenität zeigt man ebenso.

Nachweis der Schiefsymmetrie. Ist $v_i + w_i = v_j + w_j$ mit $i < j$, so folgt aus der Direkte Summe- Eigenschaft $v_i = v_j$ und $w_i = w_j$. In der Summe (39) verschwinden also alle Terme, bei denen v_i und v_j beide in ϕ als Argument stehen. Übrig bleibt

$$\begin{aligned} &(-1)^i \phi(v_0, \ldots, \widehat{v_i}, \ldots, v_j, \ldots, v_n)\psi(w_i) + (-1)^j \phi(v_0, \ldots, v_i, \ldots, \widehat{v_j}, \ldots, v_n)\psi(w_j) \\ &\quad = (-1)^i \left(\phi(v_0, \ldots, \widehat{v_i}, \ldots, v_j, \ldots, v_n) + (-1)^{i+j} \phi(v_0, \ldots, v_i, \ldots, \widehat{v_j}, \ldots, v_n)\right) \psi(w_i). \end{aligned} \tag{40}$$

Beachten Sie, daß $v_i = v_j$. Vertauscht man im zweiten Summanden rechts das v_i der Reihe nach mit $v_{i+1}, \ldots, v_{j-1}$, so landet es auf dem Platz von v_j und man erbt von den $j - i - 1$ Vertauschungen ein Vorzeichen $(-1)^{j-i+1} = (-1)(-1)^{i+j}$. Also verschwindet auch (40).

Schließlich ist $\omega \neq 0$. Wählt man nämlich eine Basis $(b_0, \ldots, b_n)$ von $\tilde{V}$, so daß $(b_1, \ldots, b_n)$ eine Basis von V und b_0 eine solche von W ist, so ist

$$\omega(b_0, \ldots, b_n) = \phi(b_1, \ldots, b_n)\psi(b_0) \neq 0.$$

□

Zusammen mit dem Korollar 10 finden wir also

Korollar 11. *Die alternierenden n-Formen auf einem n-dimensionalen K-Vektorraum bilden einen 1-dimensionalen Vektorraum $\Lambda^n V^*$, bestehend aus der 0-Form und den Determinantenformen. Jede Determinantenform ist eine Basis von $\Lambda^n V^*$.*

Bemerkung. Auch die alternierenden k-Formen eines n-dimensionalen Vektorraums bilden einen K-Vektorraum $\Lambda^k V^*$. Dieser hat die Dimension $\binom{n}{k}$. Die Theorie der k-Formen bezeichnet man als *Multilineare Algebra*. Wir kommen vielleicht im nächsten Semester darauf zurück.

Definition 46 (Orientierung). Sei V ein n-dimensionaler $\mathbb{R}$-Vektorraum, $1 \leq n < \infty$. (Statt $\mathbb{R}$ könnte man einen angeordneten Körper nehmen, aber diese Verallgemeinerung spielt wohl keine Rolle.) Zwei Basen $(v_1, \ldots, v_n)$ und $(w_1, \ldots, w_n)$ von V heißen *gleichorientiert*, wenn für eine (und dann alle) Determinantenformen ω die reellen Zahlen $\omega(v_1, \ldots, v_n)$ und $\omega(w_1, \ldots, w_n)$ gleiches Vorzeichen haben. „Gleichorientiert" ist eine Äquivalenzrelation auf der Menge aller Basen von V. Jede Äquivalenzklasse heißt *eine Orientierung von* V. Ein endlich-dimensionaler $\mathbb{R}$-Vektorraum zusammen mit einer Orientierung heißt ein *orientierter Vektorraum*. Es gibt auf jedem V genau 2 Orientierungen: Repräsentiert $(v_1, \ldots, v_n)$ eine Orientierung, so repräsentiert $(-v_1, \ldots, v_n)$ die andere. Bemerkung: Durch die Wahl einer Determinantenform ω wird eine Orientierung bestimmt: Die Menge aller Basen, auf denen ω positiv ist.

Sie finden Determinantenformen in den Lehrbüchern der Linearen Algebra nicht sehr häufig. Aber sie finden die Determinante von Matrizen und Endomorphismen, die wir nun definieren.

Definition 47 (Determinante eines Endomorphismus). Seien V ein n-dimensionaler K-Vektorraum und $f \in \mathrm{End}(V)$ ein Endomorphismus von V. Wähle eine Determiantenform ω auf V. Dann ist $f^*\omega$ mit

$$f^*\omega(v_1, \ldots, v_n) := \omega(f(v_1), \ldots, f(v_n))$$

eine alternierende n-Form auf V. Also gibt es ein eindeutig bestimmtes $\lambda \in K$ mit $f^*\omega = \lambda\omega$. Dieses λ ist unabhängig von der gewählten Determinantenform ω und heißt *die Determinante* von f, bezeichnet als $\det f$. Also gilt

$$f^*\omega = (\det f)\omega.$$

Beweis der Wohldefiniertheit. Daß $f^*\omega \in \Lambda^n V^*$ ist offensichtlich. Wegen $f^*(\mu\omega) = \mu(f^*\omega)$ ist aber auch die Unabhängigkeit vom gewählten ω nach Korollar 11 klar. □

Bemerkung. Nun haben wir schon drei Bedeutungen von f^*: Duale Abbildung, adjungierte Abbildung und die Abbildung auf den alternierenden n-Formen auf einem n-dimensionalen Vektorraum. Um welche Bedeutung es sich handelt wird deutlich, wenn man angibt, von wo nach wo die Abbildung geht, oder in welchem Raum das Argument liegt. Ist $f \in \mathrm{End}(V)$, so ist

1. $f^* : V^* \to V^*$ ist die duale Abbildung
2. $f^* : V \to V$ ist die adjungierte Abbildung. Die existiert allerdings nur, wenn V euklidisch oder unitär ist.
3. $f^* : \Lambda^n V^* \to \Lambda^n V^*$ ist die oben definierte Abbildung.

Kritisch ist der Fall $n = 1$, denn $\Lambda^1 V^* = V^*$. Glücklicherweise stimmen dann aber die Bedeutungen 1 und 3 überein:

$$f^*\omega(v) = \omega(f(v)).$$

Lemma 12. *Ein Endomorphismus $f \in \mathrm{End}(V)$ eines n-dimensionalen Vektorraumes V ist genau dann ein Isomorphismus, wenn* $\det f \neq 0$.

Beweis. Seien $(v_1, \dots, v_n)$ eine Basis von V und ω eine Determinantenform. Ist f ein Isomorphismus, so ist $(f(v_1), \dots, f(v_n))$ wieder eine Basis und daher $\omega(f(v_1), \dots, f(v_n)) \neq 0$. Ist andrerseits f kein Isomorphismus, so sind $(f(v_1), \dots, f(v_n))$ linear abhängig und daher ist $\omega(f(v_1), \dots, f(v_n)) = 0$. □

Satz 37 (Determinanten-Multiplikationssatz). *Sind V ein endlich-dimensionaler K-Vektorraum und $f, g \in \mathrm{End}(V)$, so gilt*

$$\det(f \circ g) = (\det f)(\det g)$$

Insbesondere ist f genau dann ein Automorphismus, wenn $\det f \neq 0$.

Beweis. Sei ω eine Determinantenform auf V und sei $(v_1, \dots, v_n)$ eine Basis von V. Dann ist

$$\begin{aligned}\det(f \circ g)\omega(v_1, \dots, v_n) &= (f \circ g)^*\omega(v_1, \dots, v_n) = \omega(f(g(v_1)), \dots, f(g(v_n)))\\ &= f^*\omega(g(v_1), \dots, g(v_n)) = (\det f)\omega(g(v_1), \dots, g(v_n))\\ &= (\det f)(\det g)\omega(v_1, \dots, v_n).\end{aligned}$$

Weil für eine Basis $\omega(v_1, \dots, v_n) \neq 0$, folgt daraus die Behauptung. □

Bedeutung der Determinante. Wenn wir unserer vorangestellten Motivation Glauben schenken und Determinantenformen als ein Maß für das Volumen des von den Argumentvektoren aufgespannten Parallelotops interpretieren, dann ist $\det f$ nach Definition gerade die *Volumenverzerrung* durch den Endomorphismus f. Diese Bedeutung der Determinante ist in der Analysis von großer Wichtigkeit (Transformationsformel für n-dimensionale Integrale). Die zweite wichtige Bedeutung der Determinante beruht auf folgender Überlegung: *Eigenvektoren* eines Endomorphismus f sind Vektoren $v \neq 0$, die durch f bis auf eine Streckung um einen Faktor λ, den sogenannten *Eigenwert*, festgelassen werden:

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

Diese Gleichung kann man auch so schreiben:

$$(f - \lambda\,\mathrm{Id})(v) = 0.$$

Gesucht sind Werte λ, für die diese Gleichung eine nicht-triviale Lösung v besitzt. Das ist aber genau dann der Fall, wenn $\mathrm{Kern}(f - \lambda\,\mathrm{Id}) \neq \{0\}$, d.h. wenn f kein Isomorphismus ist. Das liefert die sogenannte charakteristische Gleichung für die Berechnung der Eigenwerte:

$$\det(f - \lambda\,\mathrm{Id}) = 0.$$

Auf diesen Themenkreis kommen wir in der Linearen Algebra noch zurück.

Definition 48 (Determinante einer Matrix). Für eine (n, n)-Matrix A über dem Köper K sei $f_A : K^n \to K^n$ die zugehörige lineare Abbildung. Dann definieren wir die Determinante von A durch

$$\boxed{\det A := \det f_A.}$$

Notation: Man schreibt auch

$$\det A = \begin{vmatrix} a_{11} & \dots & a_{nn} \\ \vdots & & \vdots \\ a_{n1} & \dots & a_{nn} \end{vmatrix}.$$

Weil man im K^n eine kanonische Basis, nämlich $(e_1, \dots, e_n)$ hat, gibt es dort auch eine kanonische Determinantenform ω_n, die charakterisiert ist durch

$$\boxed{\omega_n(e_1, \dots, e_n) = 1.}$$

Beispiel 60. Auf dem K^2 ist mit $x = \sum x_i e_i, y = \sum y_i e_i$

$$\omega_2(x, y) = x_1 y_2 \omega_2(e_1, e_2) + x_2 y_1 \omega_2(e_2, e_1) = x_1 y_2 - x_2 y_1.$$

Auf dem K^3 liefert eine ähnliche Rechnung

$$\begin{aligned} \omega_3(x, y, z) = {} & x_1 y_2 z_3 + x_2 y_3 z_1 + x_3 y_1 z_2 \\ & - x_2 y_1 z_3 - x_3 y_2 z_1 - x_1 y_3 z_2. \end{aligned}$$

$\square$

Benutzt man nun für die Definition der Determinante einer (n, n)-Matrix die kanonische Determiantenform, so findet man

$$\det A = (\det f_A) \underbrace{\omega_n(e_1, \dots, e_n)}_{=1} = \omega_n(f_A(e_1), \dots, f_A(e_n)) = \omega_n(a_1, \dots, a_n).$$

Dabei haben wir benutzt, daß die Spalten der Matrix A, die wir hier mit a_j bezeichnen, gerade die Bilder $f_A(e_j)$ der kanonischen Basisvektoren sind. Wir formulieren das als

Satz 38 (Determinante einer Matrix). *Sei ω_n die kanonische Volumenform auf K^n. Dann ist die Determinante der (n, n)-Matrix*

$$A = (a_1, \dots, a_n)$$

mit den Spalten a_j gegeben durch

$$\det A = \omega_n(a_1, \dots, a_n).$$

Auf diese Weise ist *die Determinante selbst eine Determiantenform auf dem K^n.*

Beispiel 61. Aus Beispiel 60 finden wir

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11} a_{22} - a_{21} a_{12},$$

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = a_{11} a_{22} a_{33} + a_{21} a_{32} a_{13} + a_{31} a_{12} a_{23} - a_{21} a_{12} a_{33} - a_{31} a_{22} a_{13} - a_{11} a_{32} a_{23}.$$

Die $(2, 2)$- bzw. $(3, 3)$-Determinanten berechnen sich also nach folgendem Schema:

$$\left| \begin{array}{ccc} a_{11} & & a_{12} \\ & \times & \\ a_{21} & & a_{22} \end{array} \right| \quad \text{bzw.} \quad \left| \begin{array}{ccccccccccc} a_{11} & & a_{12} & & a_{13} & & \vdots & & a_{11} & & a_{12} \\ & \searrow & & \times & & & \times & & & \nearrow & \\ a_{21} & & a_{22} & & a_{23} & & \vdots & & a_{21} & & a_{22} \\ & \times & & \times & & & \times & & & \searrow & \\ a_{31} & & a_{32} & & a_{33} & & \vdots & & a_{31} & & a_{32} \end{array} \right|.$$

Dabei sind die „Abwärts-Diagonalen“ positiv, die „Aufwärts-Diagonalen“ negativ zu nehmen. Für $(3,3)$-Matrizen nennt man dies auch die Sarrus-Regel (Jägerzaun-Regel).

$\square$

Satz 39 (Determinante der Darstellungsmatrix). *Seien V ein K-Vektorraum mit Basis $\mathbf{v} = (v_1, \ldots, v_n)$ und $f \in \operatorname{End}(V)$. Ist $A = A_{\mathbf{v}}^{\mathbf{v}}(f)$ die Darstellungsmatix von f, so gilt*

$$\det A = \det f.$$

Beachten Sie: Für $V = K^n$ mit der kanonischen Basis war das gerade die Definition der Determinante einer Matrix.

Beweis des Satzes. Sei ω die Determinantenform auf V, die auf $\mathbf{v}$ den Wert 1 hat. Sei

$$f(v_j) = \sum_i a_{ij} v_i.$$

Dann gilt

$$\begin{aligned}
\det f &= (\det f)\omega(v_1, \ldots, v_n) \\
&= \omega(f(v_1), \ldots, f(v_n)) \\
&= \omega(\sum_{i_1} a_{i_1 1} v_{i_1}, \ldots, \sum_{i_n} a_{i_n n} v_{i_n}) \\
&= \sum_{i_1} \ldots \sum_{i_n} a_{i_1 1} \ldots a_{i_n n}\, \omega(v_{i_1}, \ldots, v_{i_n}) \\
&= \sum_{i_1} \ldots \sum_{i_n} a_{i_1 1} \ldots a_{i_n n}\, \omega_n(e_{i_1}, \ldots, e_{i_n}) \\
&= \omega_n(\sum_{i_1} a_{i_1 1} e_{i_1}, \ldots, \sum_{i_n} a_{i_n n} e_{i_n}) \\
&= \omega_n(f_A(e_1), \ldots, f_A(e_n)) \\
&= \det A.
\end{aligned}$$

$\square$

15 Berechnung von Determinanten

Mit Satz 39 ist das Problem der effektiven Determinantenberechnung für Endomorphismen zurückgeführt auf das Problem der Determinantenberechnung für Matrizen. Für „kleine“ Matrizen ist letzteres im Beispiel 61 gelöst. Wir wollen es nun auch für größere Matrizen betrachten.

In diesem Abschnitt sei K stets ein Körper und ω_n die kanonische Determinantenform auf K^n.

Definition 49 (Permutationen). Eine bijektive Abbildung $\sigma : \{1, \ldots, n\} \to \{1, \ldots, n\}$ nennt man eine Permutation von $\{1, \ldots, n\}$. Die Menge der Permutationen von $\{1, \ldots, n\}$ bildet eine Gruppe $\mathcal{S}_n$, und man kann durch vollständige Induktion zeigen, daß sie $n!$ Elemente hat. Für $\sigma \in \mathcal{S}_n$ ist

$$\operatorname{sign} \sigma = \omega_n(e_{\sigma(1)}, \ldots, e_{\sigma(n)})$$

nach Definition von ω_n gerade $+1$ oder -1. Man nennt diese Zahl das *Signum* oder das *Vorzeichen* der Permutation σ.

Bemerkung. Es ist $\operatorname{sign}\sigma = +1$ genau dann, wenn man $(\sigma(1), \ldots, \sigma(n))$ durch eine *gerade* Anzahl von Vertauschungen in die Reihenfolge $(1, \ldots, n)$ bringen kann. (Beachten Sie, daß die Anzahl von Vertauschungen keineswegs eindeutig bestimmt ist, wohl aber ihre Parität.)

Aus der Bemerkung folgt

Lemma 13. *Sei V ein n-dimensionaler K-Vektorraum. Dann gilt für $\omega \in \Lambda^n V^*$ und für $\sigma, \tau \in \mathcal{S}_n$*

$$\begin{aligned}\omega(v_{\sigma(1)}, \ldots, v_{\sigma(n)}) &= \operatorname{sign}\sigma\, \omega(v_1, \ldots, v_n),\\ \operatorname{sign}(\sigma \circ \tau) &= \operatorname{sign}\sigma \operatorname{sign}\tau,\\ \operatorname{sign}(\sigma^{-1}) &= \operatorname{sign}\sigma.\end{aligned}$$

Beweis. Die erste Formel folgt direkt aus der Bemerkung. Dann ist aber

$$\begin{aligned}\operatorname{sign}(\sigma \circ \tau) = \omega_n(e_{\sigma(\tau(1))}, \ldots, e_{\sigma(\tau(n))}) &= \operatorname{sign}\sigma\, \omega_n(e_{\tau(1)}, \ldots, e_{\tau(n)})\\ &= \operatorname{sign}\sigma \operatorname{sign}\tau\, \omega_n(e_1, \ldots, e_n).\end{aligned}$$

Das beweist die zweite Formel. Aus dieser folgt die dritte, weil

$$\operatorname{sign}\sigma \operatorname{sign}\sigma^{-1} = \operatorname{sign}(\sigma\sigma^{-1}) = \operatorname{sign}\mathrm{Id} = +1.$$

□

Satz 40 (Leibniz-Formel). *Für jede (n,n)-Matrix $A = (a_{ij})$ ist*

$$\det A = \sum_{\sigma \in \mathcal{S}_n} \operatorname{sign}\sigma\, a_{\sigma(1)1} \ldots a_{\sigma(n)n}.$$

Die Determinante erhält man also, indem man alle Produkte von Matrixkomponenten bildet, die aus jeder Zeile und jeder Spalten genau ein Element enthalten, und diese mit „geeigneten" Vorzeichen aufsummiert.

Beweis. Es ist

$$\begin{aligned}\det A &= \omega_n\Big(\sum_{i_1} a_{i_1 1} e_{i_1}, \ldots, \sum_{i_n} a_{i_n n} e_{i_n}\Big)\\ &= \sum_{i_1} \ldots \sum_{i_n} a_{i_1 1} \ldots a_{i_n n}\, \omega(e_{i_1}, \ldots, e_{i_n}).\end{aligned}$$

Dabei braucht man nur die Summanden zu berücksichtigen, bei denen die $(e_{i_1}, \ldots, e_{i_n})$ paarweise verschieden sind, d.h. man muß nur über die Permuationen summieren:

$$\det A = \sum_{\sigma \in \mathcal{S}_n} a_{\sigma(1)1} \ldots a_{\sigma(n)n} \operatorname{sign}\sigma.$$

□

Damit ist das Problem der Determinantenberechnung im Prinzip gelöst. Allerdings hat man nach dieser Formel $n!$ Produkte zu bilden und zu summieren; für größere n sind das sehr viele. Wir sind daher noch an anderen Methoden interessiert, die es zum Beispiel erlauben, spezielle Eigenschaften von A auszunutzen.

Korollar 12. *Für die Transponierte A^T einer (n,n)-Matrix gilt*

$$\boxed{\det A^T = \det A.}$$

Insbesondere ist die Determinante eine Matrix auch multilinear in den Zeilen.

Beweis. Aus der Leibniz-Formel und der Definiton der Transponierten folgt

$$\det A^T = \sum_{\sigma \in \mathcal{S}_n} \operatorname{sign} \sigma \, a_{1\sigma(1)} \dots a_{n\sigma(n)}$$

Ordnet man dies Produkte jeweils nach dem zweiten Index um, so ist werden sie zu

$$a_{\sigma^{-1}(1)1} \dots a_{\sigma^{-1}(n)n}.$$

Also ist

$$\begin{aligned} \det A^T &= \sum_{\sigma \in \mathcal{S}_n} \operatorname{sign} \sigma \, a_{\sigma^{-1}(1)\,1} \dots a_{\sigma^{-1}(n)\,n} = \sum_{\sigma \in \mathcal{S}_n} \operatorname{sign} \sigma^{-1} \, a_{\sigma^{-1}(1)1} \dots a_{\sigma^{-1}(n)n} \\ &= \sum_{\tau \in \mathcal{S}_n} \operatorname{sign} \tau \, a_{\tau(1)1} \dots a_{\tau(n)n} = \det A. \end{aligned}$$

□

Bemerkung. Aus dem letzten Korollar ergibt sich, daß

$$\boxed{\det f^* = \det f}$$

für die duale Abbildung $f^* : V^* \to V^*$ zu einem Endomorphismus $f \in \operatorname{End}(V)$. Gibt es dafür einen direkten Beweis ohne den Umweg über Darstellungsmatrizen?

Lemma 14. *Seien $i, j \in \{1, \dots, n\}$. Wir definieren eine lineare Abbildung*

$$\phi_i : K^{n-1} \to K^n$$

durch

$$\phi_i \begin{pmatrix} x_1 \\ \vdots \\ x_{n-1} \end{pmatrix} := \begin{pmatrix} x_1 \\ \vdots \\ x_{i-1} \\ 0 \\ x_i \\ x_{n-1} \end{pmatrix}.$$

Es ist also

$$\phi_i(e_j) = \begin{cases} e_j & \textit{für } j < i, \\ e_{j+1} & \textit{für } j \geq i, \end{cases}$$

Beachten Sie, daß die e's auf der linken Seite die Basisvektoren im K^{n-1}, auf der rechten Seite aber die im K^n sind.

Wir definieren weiter eine $(n-1)$-Form ω_{ij} auf K^{n-1} durch

$$\begin{aligned} \omega_{ij}(v_1, \dots, v_{n-1}) &:= \omega_n(\phi_i(v_1), \dots, \phi_i(v_{j-1}), e_i, \phi_i(v_j), \dots, \phi_i(v_{n-1})) \\ &= \begin{vmatrix} v_{11} & \dots & v_{1\,j-1} & 0 & v_{1j} & \dots & v_{1\,n-1} \\ \vdots & & & \vdots & & & \vdots \\ v_{i-1\,1} & \dots & v_{i-1\,j-1} & 0 & v_{i-1\,j} & \dots & v_{i-1\,n-1} \\ 0 & \dots & 0 & 1 & 0 & \dots & 0 \\ v_{i\,1} & \dots & v_{i\,j-1} & 0 & v_{i\,j} & \dots & v_{i\,n-1} \\ \vdots & & & \vdots & & & \vdots \\ v_{n-1\,1} & \dots & v_{n-1\,j-1} & 0 & v_{n-1\,j} & \dots & v_{n-1\,n-1} \end{vmatrix} \end{aligned}$$

Dann gilt

$$\omega_{ij} = (-1)^{i+j}\omega_{n-1}.$$

Dabei bezeichnet ω_{n-1} die kanonische Determinantenform auf K^{n-1}.

Beweis. Es ist klar, daß ω_{ij} multilinear und alternierend ist. Weiter ist

$$\begin{aligned}\omega_{ij}(v_1,\ldots,v_{n-1}) &= \omega_n(\phi_i(v_1),\ldots,\phi_i(v_{j-1}),e_i,\phi_i(v_j),\ldots,\phi_i(v_{n-1}))\\ &= (-1)^{n-1-j+1}\omega_n(\phi_i(v_1),\ldots,\phi_i(v_{n-1}),e_i)\\ &= (-1)^{n-j}\omega_n(\phi_i(v_1),\ldots,\phi_i(v_{n-1}),e_i).\end{aligned}$$

Daraus folgt

$$\begin{aligned}\omega_{ij}(e_1,\ldots,e_{n-1}) &= (-1)^{n-j}\omega_n(\phi_i(e_1),\ldots,\phi_i(e_{n-1}),e_i)\\ &= (-1)^{n-j}\omega_n(e_1,\ldots,e_{i-1},e_{i+1},\ldots e_n,e_i)\\ &= (-1)^{n-j}(-1)^{n-i}\omega_n(e_1,\ldots,e_n)\\ &= (-1)^{i+j}\underbrace{\omega_n(e_1,\ldots,e_n)}_{=1}\\ &= (-1)^{i+j}\omega_{n-1}(e_1,\ldots,e_{n-1}).\end{aligned}$$

□

Satz 41 (Laplacescher Entwicklungssatz). *Seine $A = (a_{ij})$ eine (n,n)-Matrix und $j \in \{1,\ldots,n\}$. Sei A_{ij} die Matrix, die man aus A durch Streichen der i-ten Zeile und j-ten Spalte erhält (Kofaktormatrix). Dann gilt:*

$$\det A = \sum_{i=1}^{n}(-1)^{i+j}a_{ij}\det A_{ij}.$$

Man nennt das die Entwicklung der Determinante nach der j-ten Spalte.
Wegen Korollar 12 gilt ein entsprechender Satz auch für Zeilen.

Beweis. Wir bezeichnen mit a_j den j-ten Spaltenvektor von A und mit $a_j^i \in K^{n-1}$ den daraus durch Streichen der i-ten Komponente erhaltenen Vektor. Dann ist mit den Bezeichnungen des Lemmas

$$a_j = \phi_i(a_j^i) + a_{ij}e_i. \tag{41}$$

Nun ist

$$\det A = \omega_n(a_1,\ldots,\overset{\overset{j}{\vee}}{\sum_i a_{ij}e_i},\ldots,a_n) = \sum_i a_{ij}\omega_n(a_1,\ldots,\overset{\overset{j}{\vee}}{e_i},\ldots,a_n).$$

Wir ersetzen die a_j-Spalten durch (41). Weil ω_n altenierend ist, kann man dabei den Term $+a_{ij}e_i$ weglassen, d.h. man kann in den a-Spalten die e_i-Komponenten weglassen:

$$\begin{aligned}\det A &= \sum_i a_{ij}\omega_n(\phi_i(a_1^i),\ldots,\overset{\overset{j}{\vee}}{e_i},\ldots,\phi_i(a_n^i)) = \sum_i a_{ij}\omega_{ij}(a_1^i,\ldots,\widehat{a_j^i},\ldots,a_n^i)\\ &= \sum_i(-1)^{i+j}a_{ij}\omega_{n-1}(a_1^i,\ldots,\widehat{a_j^i},\ldots,a_n^i) = \sum_i(-1)^{i+j}a_{ij}\det A_{ij}.\end{aligned}$$

□

Beispiel 62 (Dreiecksmatrizen). Mit vollständiger Induktion und Entwicklung nach der ersten Spalte beweist man (im Kopf)

$$\begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ 0 & a_{22} & \dots & a_{2n} \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \dots & a_{nn} \end{vmatrix} = a_{11} \cdot \ldots \cdot a_{nn}.$$

Die Determinante einer Dreiecksmatrix ist das Produkt der Diagonalelemente. Das gilt auch für untere Dreiecksmatrizen.

□

Korollar 13. *Seien $A \in M(m,m;K), B \in M(n,n;K)$ und $C \in M(m,n;K)$. Dann gilt für die daraus gebildete $(m+n,m+n)$-Blockmatrix*

$$\det \begin{pmatrix} A & C \\ 0 & B \end{pmatrix} = \det A \det B.$$

Beweis. Durch vollständige Induktion über m. Für $m = 1$ folgt die Behauptung direkt aus dem Entwicklungssatz.

$(m-1) \to m$. Sei die Behauptung für $m-1$ bereits bewiesen. Ist X die Blockmatrix, so gilt für $1 \leq i, j \leq m$

$$X_{ij} = \begin{pmatrix} A_{ij} & C \\ 0 & B \end{pmatrix}$$

Durch Entwicklung nach der ersten Spalte ergibt sich also

$$\begin{aligned} \det X &= \sum_{i=1}^{m+n} (-1)^{i+1} x_{i1} \det X_{i1} \\ &= \sum_{i=1}^{m} (-1)^{i+1} x_{i1} \det X_{i1} \\ &= \sum_{i=1}^{m} (-1)^{i+1} a_{i1} \det \begin{pmatrix} A_{i1} & C \\ 0 & B \end{pmatrix} \\ &\underset{Ind.-Vor.}{=} \sum_{i=1}^{m} (-1)^{i+1} a_{i1} \det A_{i1} \det B \\ &= \left(\sum_{i=1}^{m} (-1)^{i+1} a_{i1} \det A_{i1} \right) \det B \\ &= \det A \det B \end{aligned}$$

□

Berechnung mit dem Gaußalgorithmus. Weil die Determinante von Matrizen in den Zeilen linear und alternierend ist, kann man zur Berechnung den Gaußalgorithmus heranziehen. Addition des Vielfachen einer Zeile zu einer anderen ändert die Determinante nicht: Zum Beispiel ist

$$\det \begin{pmatrix} a_1 + a_2 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} = \det \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} + \det \begin{pmatrix} a_2 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} = \det \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix}.$$

Die Vertauschung von Zeilen hingegen ändert nur das Vorzeichen, darüber muß man „Buch führen". Dieses Verfahren zur Berechnung von Determinanten ist in vielen Fällen das einfachste.

Beispiel 63 (Vandermondesche Determinante). Seien $\lambda_1, \ldots, \lambda_n \in K$. Dann gilt

$$\Delta(\lambda_1, \ldots, \lambda_n) := \det \begin{pmatrix} 1 & 1 & \ldots & 1 \\ \lambda_1 & \lambda_2 & \ldots & \lambda_n \\ \lambda_1^2 & \lambda_2^2 & \ldots & \lambda_n^2 \\ \vdots & \vdots & & \vdots \\ \lambda_1^{n-2} & \lambda_2^{n-2} & \ldots & \lambda_n^{n-2} \\ \lambda_1^{n-1} & \lambda_2^{n-1} & \ldots & \lambda_n^{n-1} \end{pmatrix} = \prod_{1 \leq i < j \leq n} (\lambda_j - \lambda_i).$$

Die Determinante verschwindet genau dann, wenn zwei der λ_i gleich sind.

Um das zu zeigen, subtrahieren wir – von unten beginnend – von jeder Zeile das λ_1-fache der vorangehenden. Dabei ändert sich die Determinante nicht. Also

$$\begin{aligned}
\Delta(\lambda_1, \ldots, \lambda_n) &= \det \begin{pmatrix} 1 & 1 & \ldots & 1 \\ 0 & \lambda_2 - \lambda_1 & \ldots & \lambda_n - \lambda_1 \\ 0 & \lambda_2^2 - \lambda_1\lambda_2 & \ldots & \lambda_n^2 - \lambda_1\lambda_n \\ \vdots & \vdots & & \vdots \\ 0 & \lambda_2^{n-2} - \lambda_1\lambda_2^{n-3} & \ldots & \lambda_n^{n-2} - \lambda_1\lambda_n^{n-3} \\ 0 & \lambda_2^{n-1} - \lambda_1\lambda_2^{n-2} & \ldots & \lambda_n^{n-1} - \lambda_1\lambda_n^{n-2} \end{pmatrix} \\
&= \det \begin{pmatrix} 1 & 1 & \ldots & 1 \\ 0 & (\lambda_2 - \lambda_1)\,1 & \ldots & (\lambda_n - \lambda_1)\,1 \\ 0 & (\lambda_2 - \lambda_1)\lambda_2 & \ldots & (\lambda_n - \lambda_1)\lambda_n \\ \vdots & \vdots & & \vdots \\ 0 & (\lambda_2 - \lambda_1)\lambda_2^{n-3} & \ldots & (\lambda_n - \lambda_1)\lambda_n^{n-3} \\ 0 & (\lambda_2 - \lambda_1)\lambda_2^{n-2} & \ldots & (\lambda_n - \lambda_1)\lambda_n^{n-2} \end{pmatrix} \\
&= \det \begin{pmatrix} (\lambda_2 - \lambda_1)\,1 & \ldots & (\lambda_n - \lambda_1)\,1 \\ (\lambda_2 - \lambda_1)\lambda_2 & \ldots & (\lambda_n - \lambda_1)\lambda_n \\ \vdots & & \vdots \\ (\lambda_2 - \lambda_1)\lambda_2^{n-3} & \ldots & (\lambda_n - \lambda_1)\lambda_n^{n-3} \\ (\lambda_2 - \lambda_1)\lambda_2^{n-2} & \ldots & (\lambda_n - \lambda_1)\lambda_n^{n-2} \end{pmatrix} \\
&= (\lambda_2 - \lambda_1)(\lambda_3 - \lambda_1) \ldots (\lambda_n - \lambda_1) \det \begin{pmatrix} 1 & \ldots & 1 \\ \lambda_2 & \ldots & \lambda_n \\ \vdots & & \vdots \\ \lambda_2^{n-3} & \ldots & \lambda_n^{n-3} \\ \lambda_2^{n-2} & \ldots & \lambda_n^{n-2} \end{pmatrix} \\
&= (\lambda_2 - \lambda_1)(\lambda_3 - \lambda_1) \ldots (\lambda_n - \lambda_1) \Delta(\lambda_2, \ldots, \lambda_n).
\end{aligned}$$

Hier haben wir neben dem Entwicklungssatz auch die Linearität in den Spalten ausgenutzt.

Durch vollständige Induktion über n ergibt sich daher für die Vandermondesche Determinante:

$$\Delta(\lambda_1, \ldots, l_n) = \prod_{1 \leq i < j \leq n} (\lambda_j - \lambda_i).$$

□

16 Lorentzsche Vektorräume

Sei V ein n-dimensionaler $\mathbb{R}$-Vektorraum, $n < \infty$.

Definition 50. Eine symmetrische Bilinearform

$$\langle ., . \rangle : V \times V \to \mathbb{R}, (v, w) \mapsto \langle v, w \rangle$$

heißt *positiv (negativ) definit*, wenn für alle $v \neq 0$

$$\langle v, v \rangle > 0 \quad (\text{bzw. } \langle v, v \rangle < 0).$$

Sie heißt *nicht-degeneriert*, wenn für alle $w \neq 0$

$$\langle ., w \rangle \neq 0.$$

Beispiel 64. Auf dem $\mathbb{R}^n$ ist das kanonische Lorentzprodukt gegeben durch

$$\langle \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \rangle := x_1 y_1 + x_{n-1} y_{n-1} - x_n y_n.$$

Es ist nicht-degeneriert, wie man aus

$$\langle \begin{pmatrix} y_1 \\ \vdots \\ -y_n \end{pmatrix}, \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \rangle = \sum y_i^2$$

ersieht.

Wir betrachten speziell den Fall $n = 2$. „Einheitsvektoren“ sind gekennzeichnet durch

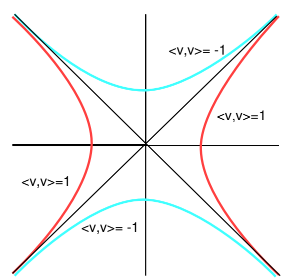
$$\langle \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \rangle = \pm 1,$$

also durch die Gleichung

$$x_1^2 - x_2^2 = 1 \quad \text{bzw.} \quad x_1^2 - x_2^2 = -1.$$

Sie liegen auf Hyperbeln. Die Vektoren $\begin{pmatrix} x \\ \pm x \end{pmatrix}$ sind sogenannte isotrope Vektoren:

$$\langle \begin{pmatrix} x \\ \pm x \end{pmatrix}, \begin{pmatrix} x \\ \pm x \end{pmatrix} \rangle = 0.$$



<v,v>=0 auf den Winkelhalbierenden

□

Lemma 15. *Sei $\langle ., . \rangle$ eine nicht-degenerierte Bilinearform auf V und sei $U \subset V$ ein Vektorunterraum, auf dem $\langle ., . \rangle$ positiv (negativ) definit ist. Dann gilt für*

$$U^\perp := \{v \in V \mid \langle v, u \rangle = 0 \text{ für alle } u \in U\},$$

daß

$$V = U \oplus U^\perp.$$

Beweis. Sei $(u_1, \ldots, u_k)$ eine ON-Basis von U, d.h.

$$\langle u_i, u_j \rangle = \pm\delta_{ij}.$$

Dabei steht das Minuszeichen, falls $\langle ., . \rangle$ auf U negativ definit ist. Dann gilt für $v \in V$:

$$v = \underbrace{\sum_{i=1}^{k} \pm\langle v, u_i \rangle u_i}_{=:P_U(v)} + (v - \sum_{i=1}^{k} \pm\langle v, u_i \rangle u_i) = P_U(v) + (v - P_U(v)).$$

Offenbar ist $P_U(V) \in U$ und für alle $u \in U$ ist

$$\langle v - P_U(v), u \rangle = \langle v, u \rangle - \sum \pm\langle v, u_i \rangle \langle u_i, u \rangle = \langle v, u \rangle - \langle v, u \rangle = 0.$$

Daher ist $V = U + U^{\perp}$. Ist $u \in U \cap U^{\perp}$, so folgt $\langle u, u \rangle = 0$, aber weil $\langle ., . \rangle$ auf U definit ist, folgt $u = 0$. Daher ist die Summe direkt. □

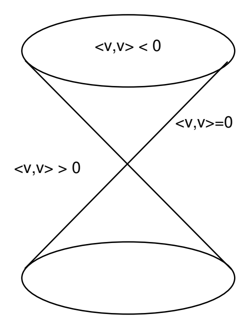
Definition 51. Sei $\langle ., . \rangle$ eine nicht-degenerierte Bilinearform auf V. Der *Positivitätsindex* von $\langle ., . \rangle$ ist die maximale Dimension eines Unterrraumes von V, auf dem $\langle ., . \rangle$ positiv definit ist. Ein reeller Vektorraum V der Dimension n, $2 \leq n < \infty$, zusammen mit einer nicht-degenerierten symmetrischen Bilinearform vom Positivitätsindex $n-1$ heißt ein *Lorentzraum.*

Beispiel 65. Der $\mathbb{R}^2$ mit dem Lorentzprodukt aus dem letzten Beispiel hat Positivitätsindex 2-1=1. Echte Unterräume sind die Geraden. Das orthogonale Komplement einer Geraden erhält man durch (euklidische) Spiegelung an einer der (euklidischen) „Winkelhalbierenden“. (Es ist egal, welche man nimmt).

Der $\mathbb{R}^n$ mit dem Lorentzprodukt hat den $(n-1)$-dimensionalen Unterraum $x_n = 0$, auf dem das Produkt positiv definit ist. Da es insgesamt aber nicht positiv definit ist, ist der Positivitätsindex $n-1$, man hat einen lorentzschen Vektorraum. Die Vektoren x mit $\langle x, x \rangle = 0$ sind charakterisiert durch die Gleichung

$$x_1^2 + \ldots + x_{n-1}^2 = x_n^2.$$

Das ist ein Doppelkegel um die x_n-Achse.



□

Definition 52. Sei $(V; \langle ., . \rangle)$ ein lorentzscher Vektorraum. Wir nennen $v \in V \backslash \{0\}$

1. *raumartig*, falls $\langle v, v \rangle > 0$,
2. *zeitartig*, falls $\langle v, v \rangle < 0$,
3. *lichtartig* oder *isotrop*, falls $\langle v, v \rangle = 0$.

Zur Kinematik der Speziellen Relativitätstheorie.

Wir betrachten einen 4-dimensionalen Lorentzraum $(V, \langle .,. \rangle)$. Die Menge der zeitartigen Vektoren zerfällt in die zwei disjunkte Mengen, die „Hälften" des Doppelkegels, so daß je zwei Vektoren aus derselben Menge negatives Skalarprodukt haben. Die Wahl einer dieser beiden Kegelhälften nennen wir eine *Zeitorientierung* und die ausgewählte Komponente die *Zukunft*. Vektoren w aus *Zukunft* nennen wir *zukunftszeitartig*. Wenn sie außerdem $\langle w, w \rangle = -1$ erfüllen, nennen wir sie ein *Inertialsystem* oder einen *kräftefreien Beobachter*. Punkte in V heißen auch *Ereignisse*.

Im folgenden sei $(V; \langle .,. \rangle)$ ein 4-dimensionaler lorentzscher Vektorraum mit einer Zeitorientierung.

1. **Relative Raum-Zeit-Zerlegung.** Ist $w \in V$ ein kräftefreier Beobachter, so ist nach dem Lemma
$$V = (\mathbb{R}w)^{\perp} \oplus \mathbb{R}w$$
und $\langle .,. \rangle$ auf $(\mathbb{R}w)^{\perp}$ positiv definit.

   Gäbe es $v \in (\mathbb{R}w)^{\perp} \setminus \{0\}$ mit
$$\langle v, v \rangle \leq 0,$$
so wäre $\langle .,. \rangle$ auf $\mathrm{Spann}(v, w)$ negativ semi-definit:
$$\langle \lambda v + \mu w, \lambda v + \mu w \rangle = \lambda^2 \langle v, v \rangle + \mu^2 \langle w, w \rangle \leq 0$$
für alle λ, μ. Andrerseits gibt es einen 3-dimensionalen Unterraum $U \subset V$, auf dem $\langle .,. \rangle$ positiv definit ist. Man verifiziert sofort, daß $U \cap \mathrm{Spann}(v, w) = \{0\}$, also $\dim V = \dim \mathrm{Spann}(v, w) + \dim U = 2 + 3 = 5$. Widerspruch.

   Wir nennen $(\mathbb{R}w)^{\perp}$ den *Raum* des Beobachters w und $\mathbb{R}w$ seine *Zeitachse*.

   Dann hat man also für $v \in V$
$$v = \underbrace{(v + \langle v, w \rangle w)}_{=:R_w(v)} \underbrace{-\langle v, w \rangle}_{=:Z_w(v)} w$$

Der kräftefreie Beobachter w kann also jedem Ereignis v • einen Ort $R_w(v)$ in seinem 3-dimensionalen euklidischen Raum $(\mathbb{R}w)^{\perp}$ und • eine Zeit $Z_w(v) = -\langle v, w \rangle \in \mathbb{R}$ zuordnen. Für $t \in \mathbb{R}$ ist $\{v \in V \mid Z_w(v) = t\}$ die Menge der Ereignisse, die für w zur Zeit t stattfinden. Das ist ein affiner 3-dimensionaler „Gleichzeitigkeitsraum", der orthogonal zu w ist.

2. **Relativität der Gleichzeitigkeit.** Seien nun $w_1, w_2 \in V$ zwei kräftefreie Beobachter. Wir setzen zur Vereinfachung
$$R_i := R_{w_i} \text{ und } Z_i := Z_{w_i}.$$

Zeitumkehr. Wir betrachten im $\mathbb{R}^4$ mit dem kanonischen Lorentzprodukt

$$w_1 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}, \quad w_2 = \begin{pmatrix} 3/4 \\ 0 \\ 0 \\ 5/4 \end{pmatrix}.$$

Offenbar ist $\langle w_1, w_1 \rangle = -1 = \langle w_2, w_2 \rangle$ und $\langle w_1, w_2 \rangle = -5/3 < 0$. Wir betrachten weiter die beiden Ereignisse

$$v = \begin{pmatrix} 10 \\ 0 \\ 0 \\ 6 \end{pmatrix} \quad \tilde{v} = \begin{pmatrix} 3 \\ 0 \\ 0 \\ 5 \end{pmatrix}.$$

Dann ist

$$\begin{aligned} Z_1(v) &= 6, \quad Z_1(\tilde{v}) = 5, \\ Z_2(v) &= 0, \quad Z_2(\tilde{v}) = -(\frac{9}{4} - \frac{25}{4}) = 4. \end{aligned}$$

Der Beobachter w_1 sieht also das Ereignis v nach $\tilde{v}$, für den Beobachter w_2 ist die Reihenfolge umgekehrt.

3. **Relativgeschwindigkeit.** Das Ereignis tw_2 für $t \in \mathbb{R}$ sieht der Beobachter w_2 für alle t an der Stelle $R_2(tw_2) = 0$. Es befindet sich in seinem System also in Ruhe. Der Beobachter w_1 sieht es zur Zeit

$$Z_1(tw_2) = -t\langle w_2, w_1 \rangle$$

an der Stelle

$$R_1(tw_2) = tR_1(w_2) = t(w_2 + \langle w_2, w_1 \rangle w_1).$$

Es bewegt sich also mit der konstanten Geschwindigkeit

$$u_{12} := \frac{R_1(w_2)}{Z_1(w_2)} \in (\mathbb{R}w_1)^\perp$$

und wir nennen das *die Geschwindigkeit von w_2 relativ zu w_1*. Wir berechnen die euklidische Länge $\|u_{12}\|$ von u_{12} in $(\mathbb{R}w_1)^\perp$.

$$\begin{aligned} \langle u_{12}, u_{12} \rangle &= \frac{1}{\langle w_2, w_1 \rangle^2} \langle w_2 + \langle w_2, w_1 \rangle w_1, w_2 + \langle w_2, w_1 \rangle w_1 \rangle \\ &= \frac{1}{\langle w_2, w_1 \rangle^2} \left(-1 - \langle w_2, w_1 \rangle^2 + 2\langle w_2, w_1 \rangle^2\right) \\ &= \frac{\langle w_2, w_1 \rangle^2 - 1}{\langle w_2, w_1 \rangle^2} < 1. \end{aligned}$$

Das liefert den

Satz von der Unerreichbarkeit der Lichtgeschwindigkeit. Die Relativgeschwindigkeit zwischen kräftefreien Beobachtern ist von Betrage < 1.

Löst man die letzte Gleichung nach $\langle w_2, w_1 \rangle$ auf und benutzt, daß dieser Ausdruck < 0 ist (beide zukunftsartig), so folgt

$$\boxed{\langle w_2, w_1 \rangle = -\frac{1}{\sqrt{1 - \|u_{12}|^2}}.} \tag{42}$$

Aus der Symmetrie der linken Seite folgt

Die Relativgeschwindigkeit von w_2 zu w_1 ist dem Betrage nach gleich der von w_1 zu w_2. Vektoriell lassen sich die beiden i.a. nicht vergleichen, da sie in den verschiedenen Räumen $(\mathbb{R}w_1)^\perp$ bzw. $(\mathbb{R}w_2)^\perp$ liegen.

4. **Zeitdilatation.** Aus der Gleichung (42) folgt für $t_2 \in \mathbb{R}$

$$\boxed{t_1 := Z_1(t_2 w_2) = \frac{t_2}{\sqrt{1 - \|u_{12}|^2}}.} \tag{43}$$

Die Ereignisse $0w_2$ und $t_2 w_2$ finden für w_2 am selben Ort 0 zu den Zeiten $Z_2(0w_2) = 0$ und

$$Z_2(t_2 w_2) = -t_2 \langle w_2, w_2 \rangle = t_2$$

statt, haben also den Zeitabstand t_2. Bezüglich w_1 haben sie den Zeitabstand $t_1 = Z_1(t_2 w_2)$, der nach (43) *größer* als t_2 ist.

5. **Längenkontraktion.** Wir betrachten nun zwei im System von w_2 ruhende Punkte $0w_2$ und $x_2 + tw_2$ mit $x_2 \in (\mathbb{R}w_2)^\perp$. Sie haben den räumlichen Abstand

$$l_2 = \|R_2(x_2)\| = \|x_2\|.$$

Der Beobachter w_1 sieht die Ereignisse $0w_2$ und $x_2 + tw_2$ gleichzeitig, wenn

$$0 = Z_1(0w_2) = Z_1(x_2 + tw_2) = -\langle x_2, w_1 \rangle - t\langle w_2, w_1 \rangle,$$

d.h. wenn

$$t = \frac{\langle x_2, w_1 \rangle}{-\langle w_2, w_1 \rangle}$$

Der räumliche Abstand l_1 der beiden Ereignisse $0w_2$ und

$$x_1 := x_2 + tw_2 = x_2 + \frac{\langle x_2, w_1 \rangle}{-\langle w_2, w_1 \rangle} w_2 \in (\mathbb{R}w_1)^\perp$$

ist, weil $x_2 \perp w_2$, gegeben durch

$$\begin{aligned}
l_1^2 :=& \|x_2\| + t^2 \langle w_2, w_2 \rangle = \|x_2\|^2 - t^2 = \|x_2\|^2 - \frac{\langle x_2, w_1 \rangle^2}{\langle w_2, w_1 \rangle^2} \\
=& \|x_2\|^2 - \frac{\langle x_2, R_2(w_1) \rangle^2}{\langle w_2, w_1 \rangle^2} = \|x_2\|^2 - \langle x_2, \underbrace{\frac{R_2(w_1)}{\langle w_2, w_1 \rangle}}_{u_{21}} \rangle^2 \\
=& \|x_2\|^2 - \|x_2\|^2 \|u_{21}\|^2 \cos\alpha,
\end{aligned}$$

wobei α der Winkel zwischen x_2 und $R_2(w_1)$ in $(\mathbb{R}w_2)^\perp$ ist. Also folgt

Im System w_2 ruhende Punkte 0 und x_2 vom Abstand l_2 haben von w_1 aus betrachtet den euklidischen Abstand

$$l_1 = l_2 \sqrt{1 - \|u_{21}\|^2 \cos^2\alpha},$$

der also i.a. kleiner als l_1 ist. α ist der Winkel zwischen der Strecke $0x_2$ und der Relativgeschwindigkeit im System w_2.

6. **Das Additionstheorem der Geschwindigkeiten.** Im folgenden betrachten wir einen dritten kräftefreien Beobachter w_0. Dann gilt

$$\begin{aligned}
\langle u_{01}, u_{02}\rangle &= \langle \frac{R_0(w_1)}{-\langle w_0, w_1\rangle}, \frac{R_0(w_2)}{-\langle w_0, w_2\rangle}\rangle \\
&= \frac{1}{\langle w_0, w_1\rangle\langle w_0, w_2\rangle}\langle w_1 + \langle w_1, w_0\rangle w_0, w_2 + \langle w_2, w_0\rangle w_0\rangle \\
&= \frac{1}{\langle w_0, w_1\rangle\langle w_0, w_2\rangle}(\langle w_1, w_2\rangle + \langle w_1, w_0\rangle\langle w_2, w_0\rangle) \\
&= -\frac{\sqrt{1-\|u_{01}\|^2}\sqrt{1-\|u_{02}\|^2}}{\sqrt{1-\|u_{12}\|^2}} + 1
\end{aligned}$$

Daraus folgt

$$1 - \|u_{12}\|^2 = \frac{(1-\|u_{01}\|^2)(1-\|u_{02}\|^2)}{(1-\langle u_{01}, u_{02}\rangle)^2}. \tag{44}$$

Sind u_{01} und u_{02} in $(\mathbb{R}w_0)^\perp$ entgegengesetzt gerichtet, so ist $\langle u_{01}, u_{02}\rangle = -\|u_{01}\|\,\|u_{02}\|$ und

$$\begin{aligned}
\|u_{12}\|^2 &= 1 - \frac{(1-\|u_{01}\|^2)(1-\|u_{02}\|^2)}{(1+\|u_{01}\|\|u_{02}\|)^2} \\
&= \frac{(\|u_{01}\| + \|u_{02}\|)^2}{(1+\|u_{01}\|\,\|u_{02}\|)^2}
\end{aligned}$$

Es folgt das

Additionstheorem der Geschwindigkeiten. $$\|u_{12}\| = \frac{\|u_{01}\| + \|u_{02}\|}{1+\|u_{01}\|\,\|u_{02}\|}.$$

7. **Das Langevinsche Zwillingsparadoxon.** Wir betrachten wieder drei kräftfreie Beobachter wie oben und nehmen an, daß für t_0, t_1, t_2 gilt

$$t_0 w_0 = t_1 w_1 + t_2 w_2.$$

Dann gilt

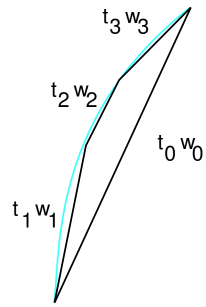
$$\begin{aligned}
t_0^2 &= -\langle t_0 w_0, t_0 w_0\rangle \\
&= -\langle t_1 w_1 + t_2 w_2, t_1 w_1 + t_2 w_2\rangle \\
&= t_1^2 + t_2^2 - 2t_1 t_2\langle w_1, w_2\rangle \\
&= t_1^2 + t_2^2 + 2t_1 t_2 \frac{1}{\sqrt{1-\|u_{12}\|^2}} \\
&\geq t_1^2 + t_2^2 + 2t_1 t_2 \\
&= (t_1 + t_2)^2
\end{aligned}$$

mit Gleichheit genau dann, wenn $\|u_{12}\| = 0$. Das ist aber äquivalent zur Bedingung

$$0 = R_1(w_2) = w_2 + \langle w_1, w_2\rangle.$$

also $w_2 = \lambda w_1$, und wegen $\langle w_1, w_1\rangle = \langle w_2, w_2\rangle$ folgt $w_1 = w_2 = w_0$.

Man kann dies verallgemeinern auf mehr als zwei Summanden auf der rechten Seite, und die rechte Seite dann als polgonale Approximation einer gekrümmten Kurve ansehen.



Langevinsches Zwillings-Paradoxon. Die Zeit „auf einer gekrümmten zukunftszeitartigen Kurve" (beschleunigter Beobachter) zwischen deren Endpunkten ist kürzer als die Zeitdifferenz zwischen diesen Endpunkten bezüglich eines kräftefreien Beobachters.