
Kapitel 2 Vektorrechnung

§ 1 Geometrische Vektoren

Inhalt:

Der physikalische Vektorbegriff, Beschreibung freier Vektoren durch Translationen, der Euklidische Raum, Geraden und Ebenen, Skalarprodukt und Vektorprodukt.

Zunächst ein paar historische Anmerkungen: Der Begriff des „Vektors“ stammt von dem Iren Sir William Rowan Hamilton (1805 - 1865), der 1843 die *Quaternionen* entdeckte, verallgemeinerte Zahlen der Gestalt $q = \alpha + ai + bj + ck$, mit $\alpha, a, b, c \in \mathbb{R}$, $i^2 = j^2 = k^2 = -1$ und $ij = -ji = k$. Den Ausdruck $ai + bj + ck$ bezeichnete Hamilton als *Vektor*. Der deutsche Gymnasiallehrer Hermann Günther Graßmann (1909 - 1877) entwickelte um 1862 die Theorie der abstrakten n -dimensionalen Vektorräume, die allerdings zunächst von der Fachwelt nicht verstanden wurde. Bis zum Ende des 19. Jahrhunderts bevorzugten die Naturwissenschaftler die Quaternionen-Theorie, erst durch Vorlesungen des Amerikaners Josiah Willard Gibbs zur „Vektoranalysis“ wurden die vektoriellen Methoden populär. Ab 1958 wurden diese Methoden nach und nach auch an den deutschen Schulen eingeführt.

Für den Physiker ist ein *Vektor* eine Größe, die durch ihren Betrag (eine positive reelle Zahl) und ihre Richtung im Raum festgelegt ist. Symbolisch wird solch ein Vektor durch einen Pfeil dargestellt. Je nach Dimension des Raumes (2, 3 oder n) kann ein Vektor dann durch zwei, drei oder n skalare Größen beschrieben werden, nämlich die Differenzen der Koordinaten von Anfangs- und Endpunkt des Pfeils. Die sind unabhängig von der Lage des Anfangspunktes.

Man benutzt Vektoren für Felder aller Art. Im Laufe der Zeit hat sich in der physikalischen Vektorrechnung eine gewisse Subkultur herausgebildet. Ein „freier Vektor“ ist ein frei beweglicher Vektor oder Pfeil, der an einem beliebigen Punkt angesetzt werden kann. Das ermöglicht es, Vektoren aneinander zu hängen („Vektoraddition“). Ein „gebundener Vektor“ ist – wie der Name sagt – an einen festen Punkt angehängt. Ein mathematisches Modell dafür könnte eine gerichtete Strecke sein. Allerdings weiß man nicht, wie man solche gerichteten Strecken addieren soll. Ein Spezialfall des gebundenen Vektors ist der „Ortsvektor“, der stets an den Nullpunkt gebunden ist. Was von all diesen verschiedenen Vektoren zu halten ist, werden wir später sehen.

Wir wollen jetzt ein mathematisches Modell für den Vektorbegriff im dreidimensionalen Raum konstruieren.

Die Punkte im Raum \mathbb{R}^3 bezeichnen wir mit den Symbolen

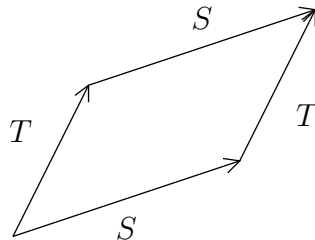
$$\mathbf{x} = (x_1, x_2, x_3), \mathbf{y} = (y_1, y_2, y_3) \text{ usw.}$$

Ein *freier Vektor* im \mathbb{R}^3 wird am besten durch eine Translation $T : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ der Gestalt $(x_1, x_2, x_3) \mapsto (x_1 + a, x_2 + b, x_3 + c)$ beschrieben. Für jeden Punkt $\mathbf{x} = (x_1, x_2, x_3) \in \mathbb{R}^3$ stellt dann das Paar $(\mathbf{x}, T(\mathbf{x}))$ einen *gebundenen Vektor* dar. Unter dem *Ortsvektor* \vec{v}_T verstehen wir den gebundenen Vektor $(\mathbf{0}, T(\mathbf{0})) = ((0, 0, 0), (a, b, c))$.

Zwei Translationen kann man miteinander verknüpfen, das Resultat ist wieder eine Translation. Diesen Vorgang bezeichnen wir als *Addition* freier Vektoren. Man kann ihn wie folgt auf die Ortsvektoren übertragen:

$$\boxed{\vec{v}_T + \vec{v}_S := \vec{v}_{T \circ S} .}$$

Führt man Translationen hintereinander aus, so hängt das Ergebnis nicht von der Reihenfolge ab. Es ist $S \circ T = T \circ S$ und $T \circ (S \circ R) = (T \circ S) \circ R$.



Außerdem kann man die Identität auch als Translation auffassen (es wird in jeder Richtung um 0 Einheiten verschoben), und natürlich ist stets $T \circ \text{id} = T$. Schließlich ist jede Translation bijektiv, und die Umkehrabbildung ist wieder eine Translation. Deshalb gilt:

1. Kommutativgesetz: $\vec{v}_T + \vec{v}_S = \vec{v}_S + \vec{v}_T$.
2. Assoziativgesetz: $(\vec{v}_T + \vec{v}_S) + \vec{v}_R = \vec{v}_T + (\vec{v}_S + \vec{v}_R)$.
3. Der *Nullvektor* $\vec{o} := \vec{v}_{\text{id}}$ erfüllt für jedes T die Gleichung

$$\vec{v}_T + \vec{o} = \vec{v}_T.$$

4. Zu jedem Vektor \vec{v}_T wird durch $-\vec{v}_T := \vec{v}_{T^{-1}}$ ein *negativer Vektor* gegeben, mit

$$\vec{v}_T + (-\vec{v}_T) = \vec{o}.$$

Diese Tatsachen sind so evident, daß wir auf einen ausführlichen Beweis verzichten können. Eine Warnung sei noch ausgesprochen: Es gibt zwar zu jedem Vektor einen negativen Vektor, aber ein Vektor allein kann niemals „positiv“ oder „negativ“ sein.

Sei jetzt λ eine reelle Zahl. Ist die Translation T durch $T(x_1, x_2, x_3) := (x_1 + a, x_2 + b, x_3 + c)$ gegeben, so definieren wir die Translation $\lambda \cdot T$ durch

$$(\lambda \cdot T)(x_1, x_2, x_3) := (x_1 + \lambda a, x_2 + \lambda b, x_3 + \lambda c).$$

Das bedeutet, daß der Pfeil, der \vec{v}_T symbolisiert, um den Faktor $|\lambda|$ gestreckt wird. Ist $\lambda < 0$, so wird zusätzlich die Richtung umgedreht. Es gilt nun:

1. $(\lambda\mu) \cdot T = \lambda \cdot (\mu \cdot T)$, für $\lambda, \mu \in \mathbb{R}$.
2. $(\lambda + \mu) \cdot T = (\lambda \cdot T) \circ (\mu \cdot T)$.
3. $\lambda \cdot (T \circ S) = (\lambda \cdot T) \circ (\lambda \cdot S)$.
4. $1 \cdot T = T$.

BEWEIS: Es sei $T(x_1, x_2, x_3) := (x_1 + a, x_2 + b, x_3 + c)$ und $S(x_1, x_2, x_3) := (x_1 + u, x_2 + v, x_3 + w)$. Die erste Aussage folgt sofort aus dem Assoziativgesetz für die Multiplikation reeller Zahlen.

2) Es ist z.B. $x_1 + (\lambda + \mu)a = (x_1 + \lambda a) + \mu a$.

3) Es ist

$$\begin{aligned} (\lambda \cdot (T \circ S))(x_1, x_2, x_3) &= (x_1 + \lambda(a + u), x_2 + \lambda(b + v), x_3 + \lambda(c + w)) \\ &= ((x_1 + \lambda a) + \lambda u, (x_2 + \lambda b) + \lambda v, (x_3 + \lambda c) + \lambda w) \\ &= ((\lambda \cdot T) \circ (\lambda \cdot S))(x_1, x_2, x_3). \end{aligned}$$

Die vierte Aussage ist wieder trivial. ■

Jetzt wird das Produkt einer reellen Zahl λ mit einem Vektor \vec{v}_T definiert durch

$$\boxed{\lambda \cdot \vec{v}_T := \vec{v}_{\lambda T}}$$

Jetzt kann man leicht die folgenden Gesetze herleiten:

1. $(\lambda\mu) \cdot \vec{v}_T = \lambda \cdot (\mu \cdot \vec{v}_T)$.
2. $(\lambda + \mu) \cdot \vec{v}_T = \lambda \cdot \vec{v}_T + \mu \cdot \vec{v}_T$.
3. $\lambda \cdot (\vec{v}_T + \vec{v}_S) = \lambda \cdot \vec{v}_T + \lambda \cdot \vec{v}_S$.
4. $1 \cdot \vec{v}_T = \vec{v}_T$.

Zum BEWEIS:

1) Es ist

$$(\lambda\mu) \cdot \vec{v}_T = \vec{v}_{(\lambda\mu) \cdot T} = \vec{v}_{\lambda \cdot (\mu \cdot T)} = \lambda \cdot \vec{v}_{\mu \cdot T} = \lambda \cdot (\mu \cdot \vec{v}_T).$$

2) Weiter ist

$$\begin{aligned} (\lambda + \mu) \cdot \vec{v}_T &= \vec{v}_{(\lambda+\mu) \cdot T} \\ &= \vec{v}_{(\lambda \cdot T) \circ (\mu \cdot T)} \\ &= \vec{v}_{\lambda \cdot T} + \vec{v}_{\mu \cdot T} \\ &= \lambda \cdot \vec{v}_T + \mu \cdot \vec{v}_T. \end{aligned}$$

3) Genauso folgt:

$$\begin{aligned} \lambda \cdot (\vec{v}_T + \vec{v}_S) &= \lambda \cdot \vec{v}_{T \circ S} \\ &= \vec{v}_{\lambda \cdot (T \circ S)} \\ &= \vec{v}_{(\lambda \cdot T) \circ (\lambda \cdot S)} \\ &= \vec{v}_{\lambda \cdot T} + \vec{v}_{\lambda \cdot S} \\ &= \lambda \cdot \vec{v}_T + \lambda \cdot \vec{v}_S. \end{aligned}$$

4) Schließlich ist $1 \cdot \vec{v}_T = \vec{v}_{1 \cdot T} = \vec{v}_T$. ■

Weil übrigens $0 \cdot T = \text{id}$ und $\lambda \cdot \text{id} = \text{id}$ ist, folgt noch:

$$0 \cdot \vec{v}_T = \vec{o} \quad \text{und} \quad \lambda \cdot \vec{o} = \vec{o}.$$

Bisher kam die Länge von Vektoren noch gar nicht vor. Der Abstand des Punktes $\mathbf{x} = (x_1, x_2, x_3)$ vom Nullpunkt ist nach Pythagoras gegeben durch

$$d(\mathbf{0}, \mathbf{x}) = \sqrt{(x_1)^2 + (x_2)^2 + (x_3)^2}.$$

Die (nicht-negative) Zahl

$$\boxed{\|\vec{v}_T\| := d(\mathbf{0}, T(\mathbf{0})) \quad (d = \text{euklidischer Abstand})}$$

nennt man die (*euklidische*) Norm von \vec{v}_T . Es gilt:

$$\|\vec{v}_T\| = 0 \iff T(\mathbf{0}) = \mathbf{0} \iff T = \text{id} \iff \vec{v}_T = \vec{o}.$$

Außerdem gilt:

1. $\|\lambda \cdot \vec{v}_T\| = |\lambda| \cdot \|\vec{v}_T\|$.
2. $\|\vec{v}_T + \vec{v}_S\| \leq \|\vec{v}_T\| + \|\vec{v}_S\|$. (Dreiecksungleichung)

Zum BEWEIS: Ist $T(x_1, x_2, x_3) = (x_1 + a, x_2 + b, x_3 + c)$, so ist $\|\vec{v}_T\| = \sqrt{a^2 + b^2 + c^2}$. Die erste Eigenschaft folgt sofort aus dieser Beziehung, denn es ist

$$\sqrt{(\lambda a)^2 + (\lambda b)^2 + (\lambda c)^2} = |\lambda| \cdot \sqrt{a^2 + b^2 + c^2}.$$

Für die zweite Eigenschaft benutzen wir die folgende aus der Elementargeometrie bekannte Tatsache: In dem aus den Ecken $\mathbf{0}$, $T(\mathbf{0})$ und $S \circ T(\mathbf{0})$ gebildeten Dreieck ist

$$\begin{aligned} \|\vec{v}_{S \circ T}\| = d(\mathbf{0}, S \circ T(\mathbf{0})) &\leq d(\mathbf{0}, T(\mathbf{0})) + d(T(\mathbf{0}), S \circ T(\mathbf{0})) \\ &= d(\mathbf{0}, T(\mathbf{0})) + d(\mathbf{0}, S(\mathbf{0})) = \|\vec{v}_T\| + \|\vec{v}_S\|. \end{aligned}$$

Ist T eine Translation, so ist der Vektor \vec{v}_T eindeutig durch die drei Komponenten des Punktes $T(\mathbf{0}) = (a, b, c)$ festgelegt. Für den Vektor \vec{v}_T verwenden wir deshalb auch die Schreibweise

$$\vec{v}_T = \begin{pmatrix} a \\ b \\ c \end{pmatrix}.$$

Dann gilt:

$$\begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} + \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} = \begin{pmatrix} a_1 + b_1 \\ a_2 + b_2 \\ a_3 + b_3 \end{pmatrix} \quad \text{und} \quad \lambda \cdot \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} = \begin{pmatrix} \lambda a_1 \\ \lambda a_2 \\ \lambda a_3 \end{pmatrix}.$$

Wir haben jetzt verschiedene Vektorbegriffe und verschiedene Beschreibungen dafür kennengelernt. Im Grunde handelt es sich aber immer um den gleichen Begriff. Egal, ob es um Translationen, Punktpaare oder Ortsvektoren geht, wir können sie alle vollständig durch drei Zahlen beschreiben. Unter dem *Euklidischen Raum* \mathbb{E}^3 verstehen wir die Menge aller dieser Vektoren. Sind $\vec{x}_1, \dots, \vec{x}_k$ Vektoren aus \mathbb{E}^3 und $\lambda_1, \dots, \lambda_k$ reelle Zahlen, so nennt man den Ausdruck

$$\lambda_1 \cdot \vec{x}_1 + \dots + \lambda_k \cdot \vec{x}_k$$

eine *Linearkombination* von $\vec{x}_1, \dots, \vec{x}_k$.

Der Raum \mathbb{R}^3 besitzt einen Nullpunkt und drei Koordinatenachsen. Auf diesen Achsen liegen die Punkte $\mathbf{e}_1 = (1, 0, 0)$, $\mathbf{e}_2 = (0, 1, 0)$ und $\mathbf{e}_3 = (0, 0, 1)$. Den Ortsvektor des Punktes \mathbf{e}_i bezeichnen wir mit \vec{e}_i . Dann ist

$$\vec{e}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad \vec{e}_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \quad \text{und} \quad \vec{e}_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}.$$

Man nennt diese drei Vektoren auch die *Einheitsvektoren*.

Ist $\vec{a} = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix}$ ein beliebiger Vektor, so ist $\vec{a} = a_1 \cdot \vec{e}_1 + a_2 \cdot \vec{e}_2 + a_3 \cdot \vec{e}_3$. Also ist

jeder Vektor aus \mathbb{E}^3 eine Linearkombination der Einheitsvektoren. Die Koeffizienten a_i sind dabei durch \vec{a} eindeutig bestimmt.

Begriffe, die man aus der Geometrie des Punktraumes \mathbb{R}^3 kennt, lassen sich oft mit Hilfe der Vektorsprache einfacher darstellen und auf höhere Dimensionen verallgemeinern.

Definition:

Es seien \vec{x}_0, \vec{v} feste Vektoren in \mathbb{E}^3 , $\vec{v} \neq \vec{o}$. Dann nennt man

$$L = \{\vec{x} \in \mathbb{E}^3 : \vec{x} = \vec{x}_0 + t \cdot \vec{v}, t \in \mathbb{R}\}$$

die *Gerade* durch \vec{x}_0 in Richtung \vec{v} . Den Vektor \vec{x}_0 bezeichnet man als *Stützvektor* der Geraden, den Vektor \vec{v} als *Richtungsvektor*.

Die Gestalt der Geraden L ist unabhängig von der Länge des Richtungsvektors \vec{v} , und als Stützvektor kann man jeden Vektor nehmen, der auf L liegt. Soll L die beiden (verschiedenen) Vektoren \vec{q}_1 und \vec{q}_2 enthalten, so muß es reelle Zahlen t_1, t_2 geben, so daß gilt:

$$\vec{q}_1 = \vec{x}_0 + t_1 \cdot \vec{v} \quad \text{und} \quad \vec{q}_2 = \vec{x}_0 + t_2 \cdot \vec{v}.$$

Also ist $\vec{q}_1 - \vec{q}_2 = (t_1 - t_2) \cdot \vec{v}$. Das bedeutet, daß man als Stützvektor den Vektor \vec{q}_1 und als Richtungsvektor den Vektor $\vec{q}_2 - \vec{q}_1$ nehmen kann. Dann ist

$$L = \{\vec{x} = \vec{q}_1 + t(\vec{q}_2 - \vec{q}_1) : t \in \mathbb{R}\}.$$

Man nennt diese Darstellung die *Zwei-Punkte-Form* von L . Für $t = 0$ erhält man \vec{q}_1 , für $t = 1$ erhält man \vec{q}_2 . Beschränkt man t auf das Einheitsintervall $[0, 1]$, so ergibt sich die Verbindungsstrecke von \vec{q}_1 und \vec{q}_2 .

Sei wieder \vec{x}_0 fest gewählt. Außerdem seien \vec{v}, \vec{w} zwei Vektoren mit der Eigenschaft, daß keiner von beiden ein Vielfaches des anderen ist (insbesondere sind dann alle beide $\neq \vec{o}$). Das bedeutet, daß sich die Geraden durch \vec{x}_0 in Richtung \vec{v} bzw. \vec{w} genau in \vec{x}_0 schneiden.

Definition:

Sind die obigen Voraussetzungen gegeben, so nennt man die Menge

$$E := \{\vec{x} = \vec{x}_0 + s \cdot \vec{v} + t \cdot \vec{w} : s, t \in \mathbb{R}\}$$

die *Ebene* durch \vec{x}_0 , die von \vec{v} und \vec{w} aufgespannt wird. \vec{x}_0 heißt *Stützvektor* und die Vektoren \vec{v} und \vec{w} heißen *Spannvektoren* der Ebene.

Satz. *Drei Vektoren, die nicht auf einer Geraden liegen, bestimmen genau eine Ebene.*

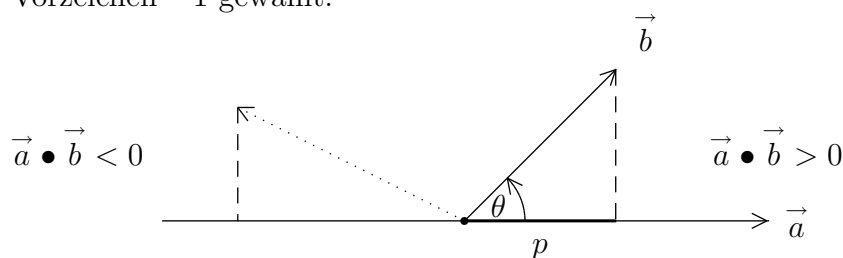
BEWEIS: Die Vektoren $\vec{q}_1, \vec{q}_2, \vec{q}_3$ mögen nicht auf einer Geraden liegen. Dann sind sie auf jeden Fall paarweise verschieden. Wir setzen $\vec{v} = \vec{q}_2 - \vec{q}_1$ und $\vec{w} = \vec{q}_3 - \vec{q}_1$. Wäre etwa $\vec{w} = \lambda \cdot \vec{v}$ (mit $\lambda \neq 0$), so würde die Gerade $L = \{\vec{x} = \vec{q}_1 + t(\vec{q}_2 - \vec{q}_1)\}$ nicht nur die Punkte \vec{q}_1 und \vec{q}_2 enthalten, sondern auch (für $t = \lambda$) den Punkt \vec{q}_3 . Das kann nicht sein. Also spannen \vec{v} und \vec{w} eine Ebene durch \vec{q}_1 auf, die dann natürlich auch \vec{q}_2 und \vec{q}_3 enthält.

Zwei Ebenen, die sich schneiden, aber nicht übereinstimmen, müssen sich entlang einer Geraden schneiden (den Beweis dafür werden wir später führen). Also kann es keine zwei Lösungen geben, denn dann lägen ja alle drei \vec{q}_i auf der Schnittgeraden.

■

Wir wollen jetzt ein neues Produkt $\vec{a} \bullet \vec{b}$ zwischen Vektoren einführen.

Dazu seien $\vec{a}, \vec{b} \in \mathbb{E}^3$. Ist einer der beiden Vektoren der Nullvektor, so soll als Produkt die Zahl 0 herauskommen. Ist $\vec{b} = \lambda \cdot \vec{a}$, so setzen wir $\vec{a} \bullet \vec{b} := \lambda \cdot \|\vec{a}\|^2$. Es bleibt der Fall, wo \vec{a} und \vec{b} eine Ebene E (durch \vec{o}) aufspannen. Diese Ebene E versehen wir so mit einem Koordinatenkreuz, daß \vec{a} in die Richtung der positiven x -Achse zeigt und \vec{b} in der oberen Halbebene liegt. Ist dann p die Länge der orthogonalen Projektion von \vec{b} auf die x -Achse, so wird $\vec{a} \bullet \vec{b}$ gleich dem Produkt aus $\|\vec{a}\|$ und p gesetzt, versehen mit einem Vorzeichen. Als Vorzeichen wird $+1$ gewählt, falls der Winkel zwischen \vec{a} und \vec{b} kleiner als $\pi/2$ ist, andernfalls wird als Vorzeichen -1 gewählt.



Steht \vec{b} auf \vec{a} senkrecht, so ist die Länge der orthogonalen Projektion gleich Null. Im Falle $0 < \theta < \pi/2$ ist offensichtlich $p = \|\vec{b}\| \cdot \cos(\theta)$. Ist $\theta > \pi/2$, so bleibt die Formel richtig, wenn man auf der linken Seite p durch $-p$ ersetzt, denn dann ist $p = \|\vec{b}\| \cdot \cos(\pi - \theta) = -\|\vec{b}\| \cdot \cos(\theta)$. Also gilt allgemein:

$$\vec{a} \bullet \vec{b} = \|\vec{a}\| \cdot \|\vec{b}\| \cdot \cos \theta,$$

wobei $\theta \in [0, \pi]$ der im Gegen-Uhrzeigersinn gemessene Winkel zwischen \vec{a} und \vec{b} ist.

Man nennt dieses Produkt das *Skalarprodukt* von \vec{a} und \vec{b} . Mit $\angle(\vec{a}, \vec{b})$ bezeichnen wir den eindeutig bestimmten Winkel $\theta \in [0, \pi]$, der durch $\cos \theta = \frac{\vec{a} \bullet \vec{b}}{\|\vec{a}\| \cdot \|\vec{b}\|}$ gegeben ist.

Rein geometrisch kann man zeigen:

$$\begin{aligned} (\vec{a} + \vec{b}) \bullet \vec{c} &= \vec{a} \bullet \vec{c} + \vec{b} \bullet \vec{c} \\ \text{und } (\lambda \cdot \vec{a}) \bullet \vec{b} &= \lambda(\vec{a} \bullet \vec{b}). \end{aligned}$$

Speziell ist $\vec{a} \bullet \vec{a} = \|\vec{a}\|^2$. Aus der geometrischen Definition des Skalarproduktes folgt nun, daß $\vec{e}_i \bullet \vec{e}_j = \delta_{ij}$ ist, für $i, j = 1, \dots, 3$. Dabei ist das *Kronecker-Symbol* δ_{ij} definiert durch

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

Satz. Sind $\vec{a} = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix}$ und $\vec{b} = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix}$ zwei Vektoren aus \mathbb{V}_3 , so ist

$$\vec{a} \bullet \vec{b} = a_1 b_1 + a_2 b_2 + a_3 b_3.$$

BEWEIS: Es ist

$$\vec{a} \bullet \vec{b} = \sum_{i,j} a_i b_j (\vec{e}_i \bullet \vec{e}_j) = \sum_{i,j} a_i b_j \delta_{ij} = \sum_{i=1}^3 a_i b_i.$$

■

Es folgen nun einige Anwendungen:

1. Die Vektoren \vec{a} und \vec{b} stehen genau dann aufeinander senkrecht, wenn

$$\vec{a} \bullet \vec{b} = 0$$

ist. Ist $\vec{a} \neq \vec{0}$ und $\vec{b} \neq \vec{0}$, so ist das genau dann der Fall, wenn $\angle(\vec{a}, \vec{b}) = \pi/2$ ist. Äquivalent ist das auch zu der Gleichung

$$\|\vec{a} + \vec{b}\|^2 = \|\vec{a}\|^2 + \|\vec{b}\|^2 \quad (\text{Satz des Pythagoras}).$$

2. Ist $\|\vec{r}\| = \|\vec{s}\|$ und $-\vec{r} + \vec{a} = \vec{s} = \vec{r} - \vec{b}$, so ist $\vec{a} \bullet \vec{b} = 0$ (Satz von Thales).

BEWEIS: Es ist $\vec{a} \bullet \vec{b} = (\vec{r} + \vec{s}) \bullet (\vec{r} - \vec{s}) = \|\vec{r}\|^2 - \|\vec{s}\|^2 = 0$.

■

3. Ist $\vec{a} = a_1 \cdot \vec{e}_1 + a_2 \cdot \vec{e}_2 + a_3 \cdot \vec{e}_3$, so ist

$$\cos \angle(\vec{a}, \vec{e}_1) = \frac{a_1}{\|\vec{a}\|}, \quad \cos \angle(\vec{a}, \vec{e}_2) = \frac{a_2}{\|\vec{a}\|} \quad \text{und} \quad \cos \angle(\vec{a}, \vec{e}_3) = \frac{a_3}{\|\vec{a}\|}.$$

Diese drei Werte werden die *Richtungskosinus(se)* von \vec{a} genannt.

Es gibt in \mathbb{E}^3 noch ein weiteres Produkt, dessen Konstruktion etwas komplizierter ist.

Definition:

Das *Vektorprodukt* $\vec{a} \times \vec{b}$ zweier Vektoren $\vec{a}, \vec{b} \in \mathbb{E}^3$ ist wieder ein Vektor aus \mathbb{E}^3 , der wie folgt definiert ist:

1. Ist $\vec{a} = \vec{b} = \vec{o}$ oder einer der beiden Vektoren \vec{a}, \vec{b} ein Vielfaches des anderen Vektors, so ist $\vec{a} \times \vec{b} = \vec{o}$.
2. Spannen \vec{a} und \vec{b} eine Ebene E auf, so ist $\vec{c} = \vec{a} \times \vec{b}$ der durch folgende Bedingungen eindeutig bestimmte Vektor:
 - (a) \vec{c} steht senkrecht auf E (ist also gleichzeitig zu \vec{a} und \vec{b} orthogonal).
 - (b) $\|\vec{c}\| = \|\vec{a}\| \cdot \|\vec{b}\| \cdot \sin \angle(\vec{a}, \vec{b})$. (Das ist möglich, weil der Winkel zwischen 0 und π liegt, der Sinus also nicht negativ wird).
 - (c) Die Vektoren \vec{a}, \vec{b} und $\vec{a} \times \vec{b}$ bilden (in dieser Reihenfolge) ein Rechtssystem. Das heißt, daß man sie (simultan, ohne ihre gegenseitige Position zueinander zu verändern) so im Raum drehen kann, daß \vec{a} in die positive x -Richtung zeigt, \vec{b} in der oberen Halbebene der x - y -Ebene liegt und $\vec{a} \times \vec{b}$ in die z -Richtung zeigt.

Der Länge von $\vec{a} \times \vec{b}$ entspricht dem Flächeninhalt des von \vec{a} und \vec{b} aufgespannten Parallelogramms. Offensichtlich ist

$$\|\vec{a} \times \vec{b}\|^2 = \|\vec{a}\|^2 \cdot \|\vec{b}\|^2 - (\vec{a} \bullet \vec{b})^2.$$

Für die Einheitsvektoren ergibt sich das Vektorprodukt besonders einfach:

$$\vec{e}_1 \times \vec{e}_2 = \vec{e}_3, \quad \vec{e}_2 \times \vec{e}_3 = \vec{e}_1 \quad \text{und} \quad \vec{e}_1 \times \vec{e}_3 = -\vec{e}_2.$$

Außerdem ist $\vec{e}_i \times \vec{e}_i = \vec{o}$ für $i = 1, 2, 3$, und es gilt:

1. $\vec{a} \times \vec{b} = -\vec{b} \times \vec{a}$,

$$2. (\lambda \vec{a}) \times \vec{b} = \lambda(\vec{a} \times \vec{b}),$$

$$3. (\vec{a}_1 + \vec{a}_2) \times \vec{b} = \vec{a}_1 \times \vec{b} + \vec{a}_2 \times \vec{b}.$$

Damit läßt sich das Vektorprodukt beliebiger Vektoren in geschlossener Form berechnen:

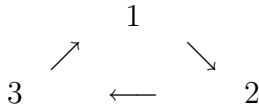
$$\begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} \times \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} = \begin{pmatrix} a_2 b_3 - a_3 b_2 \\ a_3 b_1 - a_1 b_3 \\ a_1 b_2 - a_2 b_1 \end{pmatrix}.$$

Zum Beweis schreibe man \vec{a} und \vec{b} als Linearkombinationen der Einheitsvektoren und benutze die obigen Rechenregeln.

Merkregel: Ist $\vec{a} \times \vec{b} = \vec{c}$, so ist

$$c_i = a_j b_k - a_k b_j, \quad \text{mit } (i, j, k) = (1, 2, 3), (2, 3, 1) \text{ oder } (3, 1, 2).$$

Man spricht von „zyklischen Vertauschungen“:



§ 2 Vektorräume

Inhalt:

Begriff des reellen Vektorraumes, Beispiele, einige elementare Regeln, Unterräume, Linearkombinationen, Erzeugendensysteme.

Lineare Unabhängigkeit, Basen, Austauschsatz, Dimension, affine Unterräume.

Skalarprodukt, Norm, Abstand, Winkel und Orthogonalität im \mathbb{R}^n .

Definition:

Unter einem (*reellen*) *Vektorraum* verstehen wir eine Menge V mit folgenden Eigenschaften:

1. Je zwei Elementen $\mathbf{x}, \mathbf{y} \in V$ ist ihre *Summe* $\mathbf{x} + \mathbf{y} \in V$ zugeordnet, und es gilt:

$$V1. \quad \mathbf{x} + (\mathbf{y} + \mathbf{z}) = (\mathbf{x} + \mathbf{y}) + \mathbf{z},$$

$$V2. \quad \mathbf{x} + \mathbf{y} = \mathbf{y} + \mathbf{x},$$

- V3. Es gibt genau ein Element $\mathbf{0} \in V$ mit

$$\mathbf{x} + \mathbf{0} = \mathbf{x},$$

- V4. Zu jedem $\mathbf{x} \in V$ gibt es genau ein Element $-\mathbf{x} \in V$ mit

$$\mathbf{x} + (-\mathbf{x}) = \mathbf{0}.$$

2. Ist $\alpha \in \mathbb{R}$ und $\mathbf{x} \in V$, so kann man das *Produkt* $\alpha\mathbf{x} \in V$ bilden. Für dieses Produkt gelten folgende Regeln:

$$V5. \quad (\alpha + \beta)\mathbf{x} = \alpha\mathbf{x} + \beta\mathbf{x},$$

$$V6. \quad \alpha(\mathbf{x} + \mathbf{y}) = \alpha\mathbf{x} + \alpha\mathbf{y},$$

$$V7. \quad \alpha(\beta\mathbf{x}) = (\alpha\beta)\mathbf{x},$$

$$V8. \quad 1\mathbf{x} = \mathbf{x}.$$

Beispiele.

1. Der \mathbb{R}^n ist ein reeller Vektorraum, mit

$$(x_1, \dots, x_n) + (y_1, \dots, y_n) := (x_1 + y_1, \dots, x_n + y_n),$$

$$\alpha(x_1, \dots, x_n) := (\alpha x_1, \dots, \alpha x_n).$$

2. Ein rechteckiges Zahlenschema

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1m} \\ a_{21} & a_{22} & \dots & a_{2m} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nm} \end{pmatrix}$$

wird als (reellwertige) *Matrix* mit n Zeilen und m Spalten bezeichnet. Die Einträge a_{11}, a_{12}, \dots nennt man die *Koeffizienten* der Matrix.

Die Menge aller dieser $(n \times m)$ -Matrizen wird mit $M_{n,m}(\mathbb{R})$ bezeichnet. Der erste Index gibt also immer die Anzahl der Zeilen, der zweite die Anzahl der Spalten wieder.

Man kann Matrizen wie folgt addieren und mit reellen Zahlen multiplizieren:

$$\begin{pmatrix} a_{11} & \dots & a_{1m} \\ \vdots & & \vdots \\ a_{n1} & \dots & a_{nm} \end{pmatrix} + \begin{pmatrix} b_{11} & \dots & b_{1m} \\ \vdots & & \vdots \\ b_{n1} & \dots & b_{nm} \end{pmatrix} := \begin{pmatrix} a_{11} + b_{11} & \dots & a_{1m} + b_{1m} \\ \vdots & & \vdots \\ a_{n1} + b_{n1} & \dots & a_{nm} + b_{nm} \end{pmatrix},$$

$$\alpha \begin{pmatrix} a_{11} & \dots & a_{1m} \\ \vdots & & \vdots \\ a_{n1} & \dots & a_{nm} \end{pmatrix} := \begin{pmatrix} \alpha a_{11} & \dots & \alpha a_{1m} \\ \vdots & & \vdots \\ \alpha a_{n1} & \dots & \alpha a_{nm} \end{pmatrix}$$

Die „Null-Matrix“ ist gegeben durch

$$\mathbf{O} := \begin{pmatrix} 0 & \dots & 0 \\ \vdots & & \vdots \\ 0 & \dots & 0 \end{pmatrix},$$

und außerdem setzt man

$$- \begin{pmatrix} a_{11} & \dots & a_{1m} \\ \vdots & & \vdots \\ a_{n1} & \dots & a_{nm} \end{pmatrix} := \begin{pmatrix} -a_{11} & \dots & -a_{1m} \\ \vdots & & \vdots \\ -a_{n1} & \dots & -a_{nm} \end{pmatrix}.$$

Auf diese Weise wird $M_{n,m}(\mathbb{R})$ zu einem reellen Vektorraum.

Ein Spezialfall ist der Raum $M_{n,1}(\mathbb{R})$ der Spaltenvektoren

$$\vec{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}.$$

3. Sei $I = [a, b] \subset \mathbb{R}$ ein abgeschlossenes Intervall und

$$\mathcal{C}^0(I, \mathbb{R}) := \{f : I \rightarrow \mathbb{R} \mid f \text{ stetig}\}.$$

Zwei Funktionen $f, g \in \mathcal{C}^0(I, \mathbb{R})$ werden addiert durch

$$(f + g)(x) := f(x) + g(x) \quad (\text{Superposition}).$$

Eine reelle Zahl α wird mit einer Funktion f multipliziert durch

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

Die „Nullfunktion“ ist die Funktion, die auf ganz I den Wert 0 annimmt. Mit f liegt auch die Funktion $-f$ mit $(-f)(x) := -f(x)$ in $\mathcal{C}^0(I, \mathbb{R})$.

Aus den Regeln für reellen Zahlen folgt, daß $\mathcal{C}^0(I, \mathbb{R})$ ein reeller Vektorraum ist. Seine Elemente zeigen nicht mehr viel Ähnlichkeit mit den im vorigen Abschnitt betrachteten geometrischen Vektoren. Dennoch gilt allgemein:

Die Elemente eines Vektorraumes nennt man *Vektoren*.

Bemerkungen.

1. Bei der Multiplikation von reellen Zahlen mit Vektoren bezeichnet man die reellen Zahlen auch als *Skalare*. Die Skalare sollten immer **von links** herangemultipliziert werden!
2. Beim Rechnen in Vektorräumen treten zweierlei Nullen auf. Die 0 in \mathbb{R} und der Nullvektor im Vektorraum V . Es ist

$$0\mathbf{x} = \mathbf{0} \text{ für } \mathbf{x} \in V, \quad \text{und } \alpha\mathbf{0} = \mathbf{0} \text{ für } \alpha \in \mathbb{R}.$$

BEWEIS: Es ist $0\mathbf{x} = (0 + 0)\mathbf{x} = 0\mathbf{x} + 0\mathbf{x}$. Weil auch $0\mathbf{x} + \mathbf{0} = 0\mathbf{x}$ ist, muß wegen der Eindeutigkeit des Nullvektors $0\mathbf{x} = \mathbf{0}$ sein.

Weiter ist $\alpha\mathbf{0} = \alpha(\mathbf{0} + \mathbf{0}) = \alpha\mathbf{0} + \alpha\mathbf{0}$. Mit dem gleichen Argument wie oben folgt, daß $\alpha\mathbf{0} = \mathbf{0}$ ist. ■

Satz

Sei $\alpha \in \mathbb{R}$ und $\mathbf{x} \in V$. Ist $\alpha\mathbf{x} = \mathbf{0}$, so ist $\alpha = 0$ oder $\mathbf{x} = \mathbf{0}$.

BEWEIS: Sei $\alpha\mathbf{x} = \mathbf{0}$. Ist $\alpha = 0$, so ist man fertig. Ist $\alpha \neq 0$, so gibt es in \mathbb{R} auch das Inverse α^{-1} , und es gilt:

$$\mathbf{x} = 1\mathbf{x} = (\alpha^{-1}\alpha)\mathbf{x} = \alpha^{-1}(\alpha\mathbf{x}) = \alpha^{-1}\mathbf{0} = \mathbf{0}.$$

■

Satz

Es ist $(-1)\mathbf{x} = -\mathbf{x}$.

BEWEIS: Wir verwenden die Beziehung

$$\mathbf{x} + (-1)\mathbf{x} = 1\mathbf{x} + (-1)\mathbf{x} = (1 + (-1))\mathbf{x} = 0\mathbf{x} = \mathbf{0}$$

und die Eindeutigkeit des Negativen. ■

Definition:

Sei V ein reeller Vektorraum. Eine Teilmenge $U \subset V$ heißt ein *Unterraum* von V , wenn gilt:

1. $\mathbf{0} \in U$.
2. $\mathbf{x}, \mathbf{y} \in U \implies \mathbf{x} + \mathbf{y} \in U$.
3. $\alpha \in \mathbb{R}, \mathbf{x} \in U \implies \alpha\mathbf{x} \in U$.

Ein Unterraum ist wieder ein reeller Vektorraum. Die Rechenregeln gelten alle in U , weil sie schon im übergeordneten Raum V gelten. Also muß man nur noch zeigen, daß mit jedem $\mathbf{x} \in U$ auch $-\mathbf{x}$ in U liegt. Das ist aber klar, wegen der Beziehung $-\mathbf{x} = (-1)\mathbf{x}$.

Beispiele.

1. Jeder Vektorraum besitzt den Unterraum $\{\mathbf{0}\}$.
2. Sei $\mathbf{v} \in V, \mathbf{v} \neq \mathbf{0}$. Dann ist die Menge

$$\mathbb{R}\mathbf{v} := \{\lambda\mathbf{v} \mid \lambda \in \mathbb{R}\}$$

aller „Vielfachen“ von \mathbf{v} ein Unterraum von V . Der BEWEIS dafür ist ganz einfach:

- (a) $\mathbf{0} = 0\mathbf{v}$.
- (b) Ist $\mathbf{x} = \lambda\mathbf{v}$ und $\mathbf{y} = \mu\mathbf{v}$, so ist $\mathbf{x} + \mathbf{y} = (\lambda + \mu)\mathbf{v}$.
- (c) Es ist $\alpha(\lambda\mathbf{v}) = (\alpha\lambda)\mathbf{v}$.

Definition:

Sind $\mathbf{x}_1, \dots, \mathbf{x}_k \in V$ und $\alpha_1, \dots, \alpha_k \in \mathbb{R}$, so nennt man

$$\sum_{i=1}^k \alpha_i \mathbf{x}_i := \alpha_1 \mathbf{x}_1 + \dots + \alpha_k \mathbf{x}_k$$

eine *Linearkombination* der Vektoren $\mathbf{x}_1, \dots, \mathbf{x}_k$. Das Ergebnis ist wieder ein Vektor in V .

Die Menge

$$\text{Span}(\mathbf{x}_1, \dots, \mathbf{x}_n) := \left\{ \sum_{i=1}^n \alpha_i \mathbf{x}_i \mid \alpha_i \in \mathbb{R} \text{ für } i = 1, \dots, n \right\}.$$

nennt man den *Spann* von $\mathbf{x}_1, \dots, \mathbf{x}_n$.

Satz

$U := \text{Span}(\mathbf{x}_1, \dots, \mathbf{x}_n)$ ist ein Unterraum von V .

BEWEIS: 1) $\mathbf{0} = 0\mathbf{x}_1 + \dots + 0\mathbf{x}_n$ liegt in U .

$$2) \quad \sum_{i=1}^n \alpha_i \mathbf{x}_i + \sum_{i=1}^n \beta_i \mathbf{x}_i = \sum_{i=1}^n (\alpha_i + \beta_i) \mathbf{x}_i.$$

$$3) \quad \lambda \cdot \sum_{i=1}^n \alpha_i \mathbf{x}_i = \sum_{i=1}^n (\lambda \alpha_i) \mathbf{x}_i. \quad \blacksquare$$

Man bezeichnet den von $\mathbf{x}_1, \dots, \mathbf{x}_n$ aufgespannten Unterraum auch mit $\langle \mathbf{x}_1, \dots, \mathbf{x}_n \rangle$ oder $\mathbb{R}\mathbf{x}_1 + \dots + \mathbb{R}\mathbf{x}_n$.

Ein absonderlicher Spezialfall ist der Fall $\text{Span}(\emptyset) = \{0\}$ (Begründung: Eine leere Summe ergibt immer die Null).

Definition:

Eine (endliche) Menge $E \subset V$ heißt *Erzeugendensystem* von V , wenn $\text{Span}(E) = V$ ist.

Beispiele.

1. Die
- Einheitsvektoren*

$$\mathbf{e}_1 := (1, 0, \dots, 0), \mathbf{e}_2 := (0, 1, 0, \dots, 0), \dots, \mathbf{e}_n := (0, \dots, 0, 1)$$

bilden ein Erzeugendensystem von \mathbb{R}^n , denn für einen beliebigen Vektor $\mathbf{x} = (x_1, \dots, x_n) \in \mathbb{R}^n$ gilt:

$$\mathbf{x} = x_1\mathbf{e}_1 + \dots + x_n\mathbf{e}_n.$$

Die Koeffizienten x_1, \dots, x_n sind dabei eindeutig bestimmt.

2. Im
- \mathbb{R}^2
- ist z.B. auch
- $\{(1, 3), (1, 0), (-1, -2)\}$
- ein Erzeugendensystem. Ist nämlich
- $\mathbf{x} = (x_1, x_2)$
- ein beliebiges Element des
- \mathbb{R}^2
- , so ist

$$x_2(1, 3) + x_1(1, 0) + x_2(-1, -2) = (x_2 + x_1 - x_2, 3x_2 - 2x_2) = (x_1, x_2).$$

Die Darstellung ist allerdings nicht eindeutig, es ist z.B. auch

$$\frac{x_2}{3}(1, 3) + (x_1 - \frac{x_2}{3})(1, 0) + 0(-1, -2) = (\frac{x_2}{3} + x_1 - \frac{x_2}{3}, x_2) = (x_1, x_2).$$

3. Im Raum
- $M_{n,m}(\mathbb{R})$
- der
- $(n \times m)$
- Matrizen sei
- E_{ij}
- diejenige Matrix, bei der in der
- i
- ten Zeile an der
- j
- ten Stelle eine 1 und sonst überall Nullen stehen. Dann bilden die
- $n \cdot m$
- Matrizen
- E_{ij}
- ,
- $i = 1, \dots, n$
- ,
- $j = 1, \dots, m$
- , ein Erzeugendensystem von
- $M_{n,m}(\mathbb{R})$
- , denn es gilt:

$$\left(a_{ij} \mid \begin{array}{l} i = 1, \dots, n \\ j = 1, \dots, m \end{array} \right) = \sum_{i,j} a_{ij} E_{ij}.$$

Wir haben gesehen, daß die Darstellung eines beliebigen Vektors als Linearkombination von Elementen eines Erzeugendensystems i.a. nicht eindeutig ist, daß sie es aber in Spezialfällen sein kann. Bevor wir diesen Sachverhalt klären können, müssen wir noch einen wichtigen Begriff einführen.

Definition:

Ein System $B = \{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ von Vektoren eines reellen Vektorraumes V heißt *linear unabhängig*, falls gilt:

$$\sum_{i=1}^n \alpha_i \mathbf{v}_i = 0 \implies \alpha_1 = \alpha_2 = \dots = \alpha_n = 0.$$

Ist B außerdem noch ein Erzeugendensystem von V , so nennt man B eine (endliche) *Basis* von V .

Ein System von Vektoren x_1, \dots, x_n ist also genau dann linear unabhängig, wenn sich der Nullvektor nur auf triviale Weise als Linearkombination der x_i darstellen läßt.

Beispiele.

1. Ist V ein beliebiger Vektorraum, $\mathbf{x} \in V$, $\mathbf{x} \neq \mathbf{0}$ und $\alpha\mathbf{x} = \mathbf{0}$ für ein $\alpha \in \mathbb{R}$, so muß $\alpha = 0$ sein. Daher ist das System $\{\mathbf{x}\}$ linear unabhängig.
2. Sei $\mathbf{x} \in V$, $\mathbf{x} \neq \mathbf{0}$ und $\mathbf{y} = \lambda\mathbf{x}$ für ein $\lambda \in \mathbb{R}$. Dann ist

$$\lambda\mathbf{x} + (-1)\mathbf{y} = \mathbf{0}, \text{ aber } (\lambda, -1) \neq (0, 0).$$

Also sind \mathbf{x} und \mathbf{y} nicht linear unabhängig. Man nennt sie *linear abhängig*.

3. Sei $\mathbf{x} = (a, b) \in \mathbb{R}^2$ und $\mathbf{y} = (-b, a) \in \mathbb{R}^2$, $(a, b) \neq (0, 0)$. Wir wollen zeigen, daß $\{\mathbf{x}, \mathbf{y}\}$ linear unabhängig ist.

Sei $\alpha\mathbf{x} + \beta\mathbf{y} = \mathbf{0}$. Dann ist

$$\begin{aligned} \alpha a - \beta b &= 0 \\ \text{und } \alpha b + \beta a &= 0. \end{aligned}$$

Ist $a = 0$, so muß $b \neq 0$ sein, und $\alpha b = \beta b = 0$, also $\alpha = \beta = 0$.

Ist $a \neq 0$, so multiplizieren wir die 1. Gleichung mit α und die 2. mit β . Das ergibt:

$$\begin{aligned} \alpha^2 a - \alpha\beta b &= 0 \\ \text{und } \alpha\beta b + \beta^2 a &= 0, \end{aligned}$$

also $(\alpha^2 + \beta^2)a = 0$. Das geht nur, wenn $\alpha = \beta = 0$ ist.

4. Die Einheitsvektoren $\mathbf{e}_1, \dots, \mathbf{e}_n \in \mathbb{R}^n$ sind linear unabhängig. Ist nämlich

$$\sum_{i=1}^n \alpha_i \mathbf{e}_i = \mathbf{0},$$

so ist $(\alpha_1, \dots, \alpha_n) = (0, \dots, 0)$, also $\alpha_1 = \alpha_2 = \dots = \alpha_n = 0$. Das bedeutet insbesondere, daß $\{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ eine Basis des \mathbb{R}^n ist.

5. Für jedes $n \in \mathbb{N}$ ist das System $\{1, x, x^2, x^3, \dots, x^n\} \subset \mathcal{C}^0(\mathbb{R}, \mathbb{R})$ linear unabhängig.

Ist nämlich $p := \sum_{i=0}^n \alpha_i x^i$ eine Linearkombination, die Null ergibt, so ist $p(x) \equiv 0$. Wäre einer der Koeffizienten $\alpha_i \neq 0$, so wäre $\text{grad}(p) \geq 0$, und p könnte nur endlich viele Nullstellen besitzen. Also muß $\alpha_0 = \alpha_1 = \dots = \alpha_n = 0$ sein.

Satz

Eine (endliche) Teilmenge $B = \{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ eines Vektorraumes V ist genau dann eine Basis von V , wenn sich jedes Element von V eindeutig als Linearkombination der \mathbf{v}_i schreiben läßt.

BEWEIS: a) Sei B eine Basis von V , $\mathbf{x} \in V$ beliebig vorgegeben. Da B insbesondere ein Erzeugendensystem ist, läßt sich \mathbf{x} als Linearkombination der \mathbf{v}_i schreiben. Gibt es zwei Darstellungen

$$\mathbf{x} = \sum_{i=1}^n \alpha_i \mathbf{v}_i = \sum_{i=1}^n \beta_i \mathbf{v}_i,$$

so ist

$$\mathbf{0} = \sum_{i=1}^n (\alpha_i - \beta_i) \mathbf{v}_i.$$

Da B linear unabhängig ist, muß $\alpha_i = \beta_i$ sein, für $i = 1, \dots, n$.

b) Ist umgekehrt jedes $\mathbf{x} \in V$ eindeutig als Linearkombination der \mathbf{v}_i darstellbar, so ist B offensichtlich ein Erzeugendensystem. Und da insbesondere auch der Nullvektor eindeutig darstellbar ist, muß B auch linear unabhängig sein. ■

Kennt man also eine Basis des Vektorraumes V , so kennt man auch V . Aus diesem Grund ist man daran interessiert, Basen zu finden.

Beispiel.

Sei $U := \{(x_1, x_2, x_3) \in \mathbb{R}^3 \mid x_1 + x_2 + x_3 = 0\}$.

1. Frage: Ist U ein Unterraum des \mathbb{R}^3 (und damit ein Vektorraum)?

Offensichtlich ist $\mathbf{0} \in U$. Für $\mathbf{x} = (x_1, x_2, x_3) \in U$ und $\mathbf{y} = (y_1, y_2, y_3) \in U$ ist $\mathbf{x} + \mathbf{y} = (x_1 + y_1, x_2 + y_2, x_3 + y_3)$, und offensichtlich ist

$$(x_1 + y_1) + (x_2 + y_2) + (x_3 + y_3) = (x_1 + x_2 + x_3) + (y_1 + y_2 + y_3) = 0 + 0 = 0.$$

Also liegt $\mathbf{x} + \mathbf{y}$ wieder in U .

Ist $\alpha \in \mathbb{R}$, so ist $\alpha \mathbf{x} = (\alpha x_1, \alpha x_2, \alpha x_3)$, und es gilt:

$$\alpha x_1 + \alpha x_2 + \alpha x_3 = \alpha(x_1 + x_2 + x_3) = 0.$$

Damit liegt auch $\alpha \cdot \mathbf{x}$ in U . Das bedeutet, daß U tatsächlich ein Unterraum ist.

2. Frage: Wie findet man eine Basis von U ?

Zwischen den Koordinaten eines Vektors $\mathbf{x} = (x_1, x_2, x_3) \in U$ besteht die Gleichung $x_3 = -x_1 - x_2$. Also ist $\mathbf{x} = (x_1, x_2, -x_1 - x_2)$. Weitere Bedingungen gibt es nicht. Jeder Vektor $\mathbf{x} \in U$ kann in der Form

$$\mathbf{x} = x_1(1, 0, -1) + x_2(0, 1, -1)$$

geschrieben werden. Und umgekehrt liegt jeder Vektor, der sich so darstellen läßt, in U , also insbesondere (wenn man $(x_1, x_2) = (1, 0)$ oder $= (0, 1)$ setzt) auch die Vektoren $\mathbf{a}_1 = (1, 0, -1)$ und $\mathbf{a}_2 = (0, 1, -1)$. Das bedeutet, daß \mathbf{a}_1 und \mathbf{a}_2 ein Erzeugendensystem von U bilden.

$B := \{\mathbf{a}_1, \mathbf{a}_2\}$ ist auch linear unabhängig: Ist nämlich $\alpha\mathbf{a}_1 + \beta\mathbf{a}_2 = \mathbf{0}$, so ist $(\alpha, \beta, -\alpha - \beta) = (0, 0, 0)$, also $\alpha = \beta = 0$. Damit ist B eine Basis.

Leider geht es nicht immer so einfach, und wir wissen im Augenblick noch nicht einmal, ob überhaupt jeder Vektorraum eine Basis besitzt.

Es soll nun gezeigt werden, daß jeder Unterraum des \mathbb{R}^n eine Basis besitzt, und daß alle Basen eines solchen Raumes gleich viel Elemente enthalten.

Wichtiges Hilfsmittel ist der Austauschsatz, der besagt: Hat man schon eine Basis, so kann man jedes beliebige linear unabhängige System durch Austausch geeigneter Elemente in die Basis einsetzen.

Austauschsatz

Sei $B = \{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ eine Basis des Vektorraumes V . Außerdem seien linear unabhängige Vektoren $\mathbf{w}_1, \dots, \mathbf{w}_k \in V$ gegeben, $1 \leq k \leq n$.

Nach geeigneter Numerierung der \mathbf{v}_i ist dann $\{\mathbf{w}_1, \dots, \mathbf{w}_k, \mathbf{v}_{k+1}, \dots, \mathbf{v}_n\}$ ebenfalls eine Basis von V .

BEWEIS: \mathbf{w}_1 ist eine Linearkombination der \mathbf{v}_i ,

$$\mathbf{w}_1 = \lambda_1\mathbf{v}_1 + \dots + \lambda_n\mathbf{v}_n.$$

Da $\mathbf{w}_1 \neq \mathbf{0}$ ist, muß eins der $\lambda_i \neq 0$ sein. Nach geeigneter Numerierung der \mathbf{v}_i ist $\lambda_1 \neq 0$, und es gilt:

$$\mathbf{v}_1 = \frac{1}{\lambda_1}(\mathbf{w}_1 - \lambda_2\mathbf{v}_2 - \dots - \lambda_n\mathbf{v}_n).$$

Das bedeutet, daß V auch von $\mathbf{w}_1, \mathbf{v}_2, \dots, \mathbf{v}_n$ erzeugt wird.

Diese Vektoren sind aber auch linear unabhängig. Ist nämlich $\alpha\mathbf{w}_1 + \beta_2\mathbf{v}_2 + \dots + \beta_n\mathbf{v}_n = \mathbf{0}$, so muß $\alpha = 0$ sein, denn sonst wäre \mathbf{w}_1 (und damit auch \mathbf{v}_1) eine

Linearkombination von $\mathbf{v}_2, \dots, \mathbf{v}_n$. Also ist $\beta_2\mathbf{v}_2 + \dots + \beta_n\mathbf{v}_n = \mathbf{0}$, und daher auch $\beta_2 = \dots = \beta_n = 0$.

Dieser Schritt kann nun mit der Basis $\{\mathbf{w}_1, \mathbf{v}_2, \dots, \mathbf{v}_n\}$ und dem Vektor \mathbf{w}_2 wiederholt werden. Dazu schreiben wir \mathbf{w}_2 als Linearkombination von $\mathbf{w}_1, \mathbf{v}_2, \dots, \mathbf{v}_n$.

$$\mathbf{w}_2 = \mu_1\mathbf{w}_1 + \mu_2\mathbf{v}_2 + \dots + \mu_n\mathbf{v}_n.$$

Weil $\mathbf{w}_1, \mathbf{w}_2$ linear unabhängig sind, können nicht alle der Koeffizienten μ_2, \dots, μ_n verschwinden. Nach geeigneter Numerierung ist $\mu_2 \neq 0$, also

$$\mathbf{v}_2 = \frac{1}{\mu_2}(\mathbf{w}_2 - \mu_1\mathbf{w}_1 - \mu_3\mathbf{v}_3 - \dots - \mu_n\mathbf{v}_n).$$

Das zeigt, daß V auch von $\mathbf{w}_1, \mathbf{w}_2, \mathbf{v}_3, \dots, \mathbf{v}_n$ erzeugt wird.

So fährt man fort, und nach endlich vielen Schritten ist der Satz bewiesen. ■

Folgerung 1. *Sind die Vektoren $\mathbf{a}_1, \dots, \mathbf{a}_n \in \mathbb{R}^n$ linear unabhängig, so bilden sie eine Basis des \mathbb{R}^n .*

BEWEIS: Man wende den Austauschsatz auf die Standardbasis $\{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ und auf das linear unabhängige System $\{\mathbf{a}_1, \dots, \mathbf{a}_n\}$ an. ■

Folgerung 2. *Sind $\{\mathbf{a}_1, \dots, \mathbf{a}_k\}$ und $\{\mathbf{b}_1, \dots, \mathbf{b}_s\}$ zwei Basen eines Vektorraumes V , so ist $k = s$.*

BEWEIS: Wenn nicht, dann können wir annehmen, daß $s > k$ ist. Aus dem Austauschsatz folgt, daß auch $\{\mathbf{b}_1, \dots, \mathbf{b}_k\}$ eine Basis von V ist. Aber dann ist $\mathbf{b}_{k+1} \in \text{Span}(\mathbf{b}_1, \dots, \mathbf{b}_k)$, und das ist ein Widerspruch. ■

Folgerung 3. *Sei $V \neq \{\mathbf{0}\}$ ein endlich erzeugter Vektorraum. Dann kann jedes linear unabhängige System in V zu einer Basis erweitert werden.*

BEWEIS: Sei $E = \{\mathbf{v}_1, \dots, \mathbf{v}_m\}$ ein Erzeugendensystem von V . Dann muß E mindestens einen Vektor $\neq \mathbf{0}$ enthalten, etwa \mathbf{v}_1 . Sei s die größte Zahl $\leq m$, so daß $\mathbf{v}_1, \dots, \mathbf{v}_s$ linear unabhängig sind. Dann bilden diese Vektoren sogar eine Basis B von V , denn jeder andere Vektor muß eine Linearkombination von ihnen sein.

Ist nun $\{\mathbf{w}_1, \dots, \mathbf{w}_k\}$ ein linear unabhängiges System in V , so kann man die \mathbf{w}_i gegen Elemente von B austauschen. Insbesondere muß $k \leq s$ sein. ■

Satz

Jeder Unterraum $U \subset \mathbb{R}^n$ ist endlich erzeugt und besitzt deshalb eine Basis. Alle Basen von U haben gleich viel Elemente.

Ist $B = \{\mathbf{v}_1, \dots, \mathbf{v}_k\} \subset U$, so sind die folgenden Aussagen über B äquivalent:

- 1. B ist eine Basis.*
- 2. B ist ein maximales linear unabhängiges System.*
- 3. B ist ein minimales Erzeugendensystem.*

BEWEIS: Ist $U = \{\mathbf{0}\}$, so liegt ein besonderer Fall vor. Als Basis wird dann die leere Menge aufgefaßt. Sei also $U \neq \{\mathbf{0}\}$.

Es gibt eine linear unabhängige Menge $\{\mathbf{a}_1, \dots, \mathbf{a}_k\} \subset U$, mit $k \geq 1$. Da diese Menge dann auch im \mathbb{R}^n linear unabhängig ist, muß $k \leq n$ sein. Wir wählen die Menge so, daß k maximal ist.

Ist jetzt $\mathbf{v} \in U$, so ist $\{\mathbf{v}, \mathbf{a}_1, \dots, \mathbf{a}_k\}$ linear abhängig. Es gibt also Koeffizienten $\lambda, \mu_1, \dots, \mu_k$, die nicht alle gleichzeitig verschwinden, so daß gilt:

$$\lambda \mathbf{v} + \sum_{i=1}^k \mu_i \mathbf{a}_i = \mathbf{0}.$$

Wäre $\lambda = 0$, so wäre auch $\mu_1 = \dots = \mu_k = 0$. Das kann nicht sein, es muß $\lambda \neq 0$ sein. Das bedeutet, daß es eine Darstellung der folgenden Art gibt:

$$\mathbf{v} = \sum_{i=1}^k \alpha_i \mathbf{a}_i.$$

Das bedeutet, daß $\{\mathbf{a}_1, \dots, \mathbf{a}_k\}$ ein Erzeugendensystem von U (und damit auch eine Basis) ist.

Daß je zwei Basen gleich viel Elemente enthalten, haben wir schon bewiesen.

Sei nun $B = \{\mathbf{v}_1, \dots, \mathbf{v}_k\} \subset U$. Ist B eine Basis, so ist B linear unabhängig. Da jedes andere Element von U aus den Elementen von B linear kombinierbar ist, ist B auch maximal.

Ist B ein maximales linear unabhängiges System, so muß jedes weitere Element von U linear abhängig von den Elementen von B sein. Damit ist B ein Erzeugendensystem von U . Da B linear unabhängig ist, muß B sogar ein minimales Erzeugendensystem sein.

Sei B ein minimales Erzeugendensystem. Wäre B nicht linear unabhängig, so könnte man noch weitere Elemente fortlassen, also muß B eine Basis sein. ■

Definition:

Ist V ein endlich erzeugter Vektorraum, so bezeichnet man die Anzahl der Elemente einer Basis von V als *Dimension* von V (in Zeichen : $\dim(V)$).

Ist V nicht endlich erzeugt, so setzt man $\dim(V) := \infty$.

Beispiele.

1. $\dim(\mathbb{R}^n) = n$, denn $\{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ ist eine Basis.
2. $\dim(M_{n,m}(\mathbb{R})) = n \cdot m$, denn die Matrizen E_{ij} bilden eine Basis.
3. $\dim(\mathcal{C}^0(\mathbb{R}, \mathbb{R})) = \infty$.
4. Ist $\mathbf{v} \in \mathbb{R}^n$, $\mathbf{v} \neq \mathbf{0}$, so ist die Gerade $\mathbb{R}\mathbf{v}$ ein 1-dimensionaler Unterraum des \mathbb{R}^n .
5. Sind $\mathbf{v}_1, \mathbf{v}_2$ zwei linear unabhängige Vektoren des \mathbb{R}^n , so ist die Ebene $\text{Span}(\mathbf{v}_1, \mathbf{v}_2)$ ein 2-dimensionaler Vektorraum.

Bemerkung. Man kann sogar zeigen, daß **jeder** Vektorraum eine Basis besitzt. Zum Beweis wird allerdings das „Auswahlaxiom“ der Mengenlehre gebraucht, ein Hilfsmittel, das nur mit großer Vorsicht angewandt werden sollte.

Die bisher betrachteten Geraden und Ebenen enthalten immer den Nullpunkt, wie es sich für Untervektorräume gehört. Etwas allgemeiner definiert man:

Definition:

Sei V ein reeller Vektorraum und $U \subset V$ ein k -dimensionaler Unterraum und $\mathbf{x}_0 \in V$ ein beliebiger Vektor. Dann nennt man die Menge

$$A = \mathbf{x}_0 + U := \{\mathbf{x} = \mathbf{x}_0 + \mathbf{v} \mid \mathbf{v} \in U\}$$

einen k -dimensionalen *affinen Unterraum* von V .

Im Falle $k = 1$ spricht man von affinen Geraden, im Falle $k = 2$ von affinen Ebenen.

Achtung! Ein affiner Unterraum ist i.a. kein Vektorraum, er ist nur „parallel“ zu einem Untervektorraum.

Satz

Eine Teilmenge $A \subset V$ ist genau dann ein affiner Unterraum, wenn es einen Unter(vektor)raum $U \subset V$ und ein Element $\mathbf{x}_0 \in A$ gibt, so daß gilt:

$$\mathbf{x} \in A \iff \mathbf{x} - \mathbf{x}_0 \in U.$$

BEWEIS: 1) Sei $A = \mathbf{x}_0 + U$ ein affiner Unterraum.

a) $\mathbf{x}_0 = \mathbf{x}_0 + \mathbf{0}$ liegt in A .

b) Liegt $\mathbf{x} = \mathbf{x}_0 + \mathbf{u} \in A$, so ist natürlich $\mathbf{x} - \mathbf{x}_0 = \mathbf{u} \in U$. Und ist umgekehrt $\mathbf{u} := \mathbf{x} - \mathbf{x}_0 \in U$, so liegt $\mathbf{x} = \mathbf{x}_0 + (\mathbf{x} - \mathbf{x}_0) = \mathbf{x}_0 + \mathbf{u}$ in A .

2) A erfülle das Kriterium.

a) Ist $\mathbf{x} \in A$ ein beliebiger Vektor und $\mathbf{u} := \mathbf{x} - \mathbf{x}_0 \in U$, so ist $\mathbf{x} = \mathbf{x}_0 + \mathbf{u} \in \mathbf{x}_0 + U$.

b) Ist umgekehrt $\mathbf{x} = \mathbf{x}_0 + \mathbf{u} \in \mathbf{x}_0 + U$, so ist $\mathbf{u} := \mathbf{x} - \mathbf{x}_0 \in U$, also $\mathbf{x} \in A$.

■

Der Rest dieses Abschnittes handelt nur noch vom \mathbb{R}^n . Nur dort können wir die Begriffe „Länge“ und „Richtung“ retten, die ursprünglich die Motivation für den Vektorbegriff lieferten.

Wir beginnen mit dem Skalarprodukt.

Definition:

Sind $\mathbf{x} = (x_1, \dots, x_n)$, $\mathbf{y} = (y_1, \dots, y_n)$ zwei Vektoren im \mathbb{R}^n , so bezeichnet man die reelle Zahl

$$\mathbf{x} \bullet \mathbf{y} := \sum_{i=1}^n x_i y_i$$

als das (*euklidische*) Skalarprodukt von \mathbf{x} und \mathbf{y} .

Für $n = 1$ ist dies das gewöhnliche Produkt reeller Zahlen. Im euklidischen Raum \mathbb{E}^3 ergibt sich das schon im vorigen Paragraphen behandelte Skalarprodukt.

Satz

1. $\mathbf{x} \bullet \mathbf{y} = \mathbf{y} \bullet \mathbf{x}$.
2. $(\mathbf{x}_1 + \mathbf{x}_2) \bullet \mathbf{y} = \mathbf{x}_1 \bullet \mathbf{y} + \mathbf{x}_2 \bullet \mathbf{y}$.
3. $(\alpha \cdot \mathbf{x}) \bullet \mathbf{y} = \alpha \cdot (\mathbf{x} \bullet \mathbf{y}) = \mathbf{x} \bullet (\alpha \cdot \mathbf{y})$.
4. $\mathbf{x} \bullet \mathbf{x} \geq 0$, und $\mathbf{x} \bullet \mathbf{x} = 0 \iff \mathbf{x} = \mathbf{0}$.

Die BEWEISE sind simpel. Zu Eigenschaft (4): $\mathbf{x} \bullet \mathbf{x} = \sum_{i=1}^n (x_i)^2$ ist offensichtlich ≥ 0 , und $= 0$ genau dann, wenn $x_1 = \dots = x_n = 0$ ist.

Definition:

Die Zahl $\|\mathbf{x}\| := \sqrt{\mathbf{x} \bullet \mathbf{x}} = \sqrt{(x_1)^2 + \dots + (x_n)^2}$ heißt (*euklidische*) Norm von \mathbf{x} .

Anschaulich ist die euklidische Norm des Vektors \mathbf{x} seine Länge, also der Abstand des Punktes \mathbf{x} vom Ursprung.

Satz

1. Es ist stets $\|\mathbf{x}\| \geq 0$, und $\|\mathbf{x}\| = 0 \iff \mathbf{x} = \mathbf{0}$.
2. $\|\alpha \cdot \mathbf{x}\| = |\alpha| \cdot \|\mathbf{x}\|$ für $\alpha \in \mathbb{R}$ und $\mathbf{x} \in \mathbb{R}^n$.
3. Es gilt die „Schwarzsche Ungleichung“:

$$|\mathbf{x} \bullet \mathbf{y}| \leq \|\mathbf{x}\| \cdot \|\mathbf{y}\|.$$

Gleichheit gilt genau dann, wenn \mathbf{x} und \mathbf{y} linear abhängig sind.

4. Es gilt die „Dreiecks-Ungleichung“:

$$\|\mathbf{x} + \mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\|.$$

BEWEIS: (1) und (2) sind trivial.

(3) Ist $\mathbf{y} = \mathbf{0}$, so ist die Aussage trivial. Ist $\mathbf{y} \neq \mathbf{0}$, so verwenden wir einen Trick:

Ist $\lambda \in \mathbb{R}$, so gilt:

$$\begin{aligned} 0 &\leq (\mathbf{x} - \lambda\mathbf{y}) \bullet (\mathbf{x} - \lambda\mathbf{y}) \\ &= \mathbf{x} \bullet \mathbf{x} - 2\lambda \cdot \mathbf{x} \bullet \mathbf{y} + \lambda^2 \mathbf{y} \bullet \mathbf{y}. \end{aligned}$$

Setzen wir $\lambda = \frac{\mathbf{x} \bullet \mathbf{y}}{\mathbf{y} \bullet \mathbf{y}}$, so ergibt sich:

$$\begin{aligned} 0 &\leq \mathbf{x} \bullet \mathbf{x} - 2 \frac{(\mathbf{x} \bullet \mathbf{y})^2}{\mathbf{y} \bullet \mathbf{y}} + \frac{(\mathbf{x} \bullet \mathbf{y})^2}{\mathbf{y} \bullet \mathbf{y}} \\ &= \mathbf{x} \bullet \mathbf{x} - \frac{(\mathbf{x} \bullet \mathbf{y})^2}{\mathbf{y} \bullet \mathbf{y}}, \end{aligned}$$

also

$$(\mathbf{x} \bullet \mathbf{y})^2 \leq (\mathbf{x} \bullet \mathbf{x}) \cdot (\mathbf{y} \bullet \mathbf{y}).$$

Daraus folgt die Schwarzsche Ungleichung.

(4) Es ist

$$\begin{aligned} \|\mathbf{x} + \mathbf{y}\|^2 &= (\mathbf{x} + \mathbf{y}) \bullet (\mathbf{x} + \mathbf{y}) \\ &= \|\mathbf{x}\|^2 + \|\mathbf{y}\|^2 + 2\mathbf{x} \bullet \mathbf{y} \\ &\leq \|\mathbf{x}\|^2 + \|\mathbf{y}\|^2 + 2|\mathbf{x} \bullet \mathbf{y}| \\ &\leq \|\mathbf{x}\|^2 + \|\mathbf{y}\|^2 + 2\|\mathbf{x}\| \cdot \|\mathbf{y}\| \\ &= (\|\mathbf{x}\| + \|\mathbf{y}\|)^2. \end{aligned}$$

■

Definition:

$d(\mathbf{x}, \mathbf{y}) := \|\mathbf{x} - \mathbf{y}\|$ heißt *Abstand* oder *Distanz* von \mathbf{x} und \mathbf{y} .

Speziell ist dann $\|\mathbf{x}\| = d(\mathbf{x}, \mathbf{0})$.

Die Distanz hat folgende Eigenschaften:

Satz

1. Es ist stets $d(\mathbf{x}, \mathbf{y}) \geq 0$, und $d(\mathbf{x}, \mathbf{y}) = 0 \iff \mathbf{x} = \mathbf{y}$.
2. $d(\mathbf{x}, \mathbf{y}) = d(\mathbf{y}, \mathbf{x})$.
3. Es gilt die Dreiecks-Ungleichung:

$$d(\mathbf{x}, \mathbf{y}) \leq d(\mathbf{x}, \mathbf{z}) + d(\mathbf{z}, \mathbf{y}).$$

BEWEIS: (Nur zu (3)): Es ist

$$\begin{aligned}\|\mathbf{x} - \mathbf{y}\| &= \|(\mathbf{x} - \mathbf{z}) + (\mathbf{z} - \mathbf{y})\| \\ &\leq \|\mathbf{x} - \mathbf{z}\| + \|\mathbf{z} - \mathbf{y}\|.\end{aligned}$$

■

Mit Hilfe des Skalarproduktes kann man auch Winkel zwischen Vektoren einführen.

Sind $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n \setminus \{\mathbf{0}\}$, so ist $\frac{|\mathbf{x} \bullet \mathbf{y}|}{\|\mathbf{x}\| \cdot \|\mathbf{y}\|} \leq 1$. (Scharzsche Ungleichung)

Aber zu jedem $t \in \mathbb{R}$ mit $|t| \leq 1$ gibt es genau ein $\varphi \in [0, \pi]$ mit $\cos(\varphi) = t$.

Definition:

Für Vektoren $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ mit $\mathbf{x} \neq \mathbf{0}$ und $\mathbf{y} \neq \mathbf{0}$ wird der *Winkel* $\angle(\mathbf{x}, \mathbf{y})$ definiert durch

$$\angle(\mathbf{x}, \mathbf{y}) := \arccos\left(\frac{\mathbf{x} \bullet \mathbf{y}}{\|\mathbf{x}\| \cdot \|\mathbf{y}\|}\right) \in [0, \pi].$$

Ist $\mathbf{y} = \lambda \cdot \mathbf{x}$ mit $\lambda > 0$, so ist $\angle(\mathbf{x}, \mathbf{y}) = 0$. Ist $\mathbf{y} = -\mathbf{x}$, so ist $\angle(\mathbf{x}, \mathbf{y}) = \pi$. Schließlich ist

$$\angle(\mathbf{x}, \mathbf{y}) = \frac{\pi}{2} \iff \mathbf{x} \bullet \mathbf{y} = 0.$$

Definition:

Die Vektoren \mathbf{x} und \mathbf{y} heißen *orthogonal* (oder *senkrecht*) zueinander (in Zeichen: $\mathbf{x} \perp \mathbf{y}$), wenn $\mathbf{x} \bullet \mathbf{y} = 0$ ist.

Man beachte: Ein Skalarprodukt kann verschwinden, obwohl keiner der Faktoren verschwindet. Im \mathbb{R}^2 ist z.B. $(1, 2) \bullet (-2, 1) = 0$.

Satz des Pythagoras

Ist $\mathbf{x} \perp \mathbf{y}$, so ist

$$\|\mathbf{x} + \mathbf{y}\|^2 = \|\mathbf{x}\|^2 + \|\mathbf{y}\|^2.$$

BEWEIS: Es ist

$$\|\mathbf{x} + \mathbf{y}\|^2 = (\mathbf{x} + \mathbf{y}) \bullet (\mathbf{x} + \mathbf{y}) = \|\mathbf{x}\|^2 + 2\mathbf{x} \bullet \mathbf{y} + \|\mathbf{y}\|^2.$$

Stehen die Vektoren aufeinander senkrecht, so folgt die Behauptung. ■

§ 3 Lineare Gleichungssysteme

Inhalt:

Lineare Abbildungen, Matrixschreibweise, Kern und Bild, Zusammenhang mit Injektivität und Surjektivität, Dimensionsformel.

Lineare Gleichungssysteme, Lösungsgesamtheit von homogenen und inhomogenen Systemen, Rang einer Matrix, Lösungskriterium.

Zeilen- und Spaltenoperationen, Rang-Erhaltungssatz, r -spezielle Matrizen und Gaußmatrizen, „Rückwärtseinsetzen“, Gaußsches Eliminationsverfahren, Bestimmung der Lösungsgesamtheit von homogenen und inhomogenen Systemen, Beispiele.

Definition:

V, W seien zwei reelle Vektorräume. Eine Abbildung $f : V \rightarrow W$ heißt *linear*, falls gilt:

1. $f(\mathbf{x} + \mathbf{y}) = f(\mathbf{x}) + f(\mathbf{y})$ für alle $\mathbf{x}, \mathbf{y} \in V$.
2. $f(\lambda \mathbf{x}) = \lambda f(\mathbf{x})$ für $\mathbf{x} \in V$ und $\lambda \in \mathbb{R}$.

Beispiele.

1. Der einfachste reelle Vektorraum ist \mathbb{R} selbst. Sei $f : \mathbb{R} \rightarrow \mathbb{R}$ eine lineare Abbildung. Dann ist $f(x) = f(x \cdot 1) = x f(1)$ für alle $x \in \mathbb{R}$. Die Konstante $a := f(1)$ hängt nur von f ab. Umgekehrt ist auch jede Funktion der Gestalt $f(x) = ax$ linear, denn dann gilt:

$$f(x + y) = a(x + y) = ax + ay = f(x) + f(y)$$

und

$$f(\lambda x) = a(\lambda x) = \lambda(ax) = \lambda f(x).$$

Eine affin-lineare Funktion $x \mapsto ax + b$ (mit $b \neq 0$) ist **nicht linear**. Genauso kann eine quadratische Funktion $x \mapsto ax^2 + bx + c$ mit $a \neq 0$ **nicht linear** sein.

2. Der nächst-einfache Fall ist eine lineare Abbildung $f : \mathbb{R}^n \rightarrow \mathbb{R}$. Da jeder Vektor $\mathbf{x} \in \mathbb{R}^n$ als Linearkombination $\mathbf{x} = \sum_{i=1}^n x_i \mathbf{e}_i$ geschrieben werden kann, ist

$$f(\mathbf{x}) = f(x_1 \mathbf{e}_1 + \cdots + x_n \mathbf{e}_n) = x_1 f(\mathbf{e}_1) + \cdots + x_n f(\mathbf{e}_n).$$

Setzen wir $a_i := f(\mathbf{e}_i)$, für $i = 1, \dots, n$, und $\mathbf{a} := (a_1, \dots, a_n)$, so folgt:

$$f(\mathbf{x}) = a_1 x_1 + \cdots + a_n x_n = \mathbf{a} \bullet \mathbf{x}.$$

Jede lineare Abbildung von \mathbb{R}^n nach \mathbb{R} muß so aussehen, und umgekehrt ist jede Abbildung dieser Art linear. Man spricht hier auch von *Linearformen*.

Ein Spezialfall sind die *Projektionen* $p_i : \mathbb{R}^n \rightarrow \mathbb{R}$, die durch $p_i(\mathbf{x}) = \mathbf{e}_i \bullet \mathbf{x}$ (also $p_i(x_1, \dots, x_n) = x_i$) definiert werden.

3. Ist $f : V \rightarrow W$ linear und $g : W \rightarrow U$ linear, so ist auch $g \circ f : V \rightarrow U$ linear. Der Nachweis ist ganz einfach.

Ist nun $f : \mathbb{R}^n \rightarrow \mathbb{R}^k$ linear, so gibt es Funktionen f_1, \dots, f_k mit

$$f(\mathbf{x}) = (f_1(\mathbf{x}), \dots, f_k(\mathbf{x})).$$

Man schreibt dann: $f = (f_1, \dots, f_k)$.

Offensichtlich ist $p_i \circ f = f_i$, und deshalb sind auch die Komponentenfunktionen f_i linear.

Ist umgekehrt $f = (f_1, \dots, f_k)$ mit linearen Funktionen f_i , so ist auch f linear. Man sieht das ganz einfach, es ist z.B.

$$\begin{aligned} f(\mathbf{x} + \mathbf{y}) &= (f_1(\mathbf{x} + \mathbf{y}), \dots, f_k(\mathbf{x} + \mathbf{y})) \\ &= (f_1(\mathbf{x}) + f_1(\mathbf{y}), \dots, f_k(\mathbf{x}) + f_k(\mathbf{y})) \\ &= (f_1(\mathbf{x}), \dots, f_k(\mathbf{x})) + (f_1(\mathbf{y}), \dots, f_k(\mathbf{y})) \\ &= f(\mathbf{x}) + f(\mathbf{y}). \end{aligned}$$

Das ermöglicht es, lineare Abbildungen $f : \mathbb{R}^n \rightarrow \mathbb{R}^k$ zu beschreiben. Zu jeder Komponentenfunktion f_i von f gibt es einen Vektor $\mathbf{a}_i = (a_{i1}, \dots, a_{in})$, so daß $f_i(\mathbf{x}) = \mathbf{a}_i \bullet \mathbf{x}$ ist.

4. Es gibt natürlich noch andere lineare Abbildungen. Ist z.B. $V = \{(x_1, x_2, x_3) \in \mathbb{R}^3 : x_1 + x_2 + x_3 = 0\}$ und $W = \{(x_1, x_2, x_3, x_4) \in \mathbb{R}^4 : x_2 = x_4 = 0\}$, so ist $f : V \rightarrow W$ mit $f(x_1, x_2, x_3) = (x_1 - x_2, 0, 2x_3, 0)$ eine lineare Abbildung. Tatsächlich ist f auf V definiert, die Werte von f liegen in W , und man rechnet leicht nach, daß die Eigenschaften einer linearen Abbildung erfüllt sind.

Ein wichtiges Hilfsmittel zur Beschreibung von linearen Abbildungen $f : \mathbb{R}^m \rightarrow \mathbb{R}^n$ stellen die Matrizen dar. Ein Vektor

$$\mathbf{x} = (x_1, \dots, x_m) = \sum_{i=1}^m x_i \mathbf{e}_i \in \mathbb{R}^m$$

ist durch seine Komponenten x_1, \dots, x_m eindeutig festgelegt. Das ist unabhängig davon, wie wir diese Komponenten hinschreiben. Deshalb verwenden wir manchmal die Zeilen- und manchmal die Spalten-Schreibweise, je nachdem, was gerade praktischer ist.

Sei $f = (f_1, \dots, f_n) : \mathbb{R}^m \rightarrow \mathbb{R}^n$ linear,

$$f_i(\mathbf{x}) = \mathbf{a}_i \bullet \mathbf{x},$$

mit Vektoren $\mathbf{a}_i = (a_{i1}, \dots, a_{im})$ für $i = 1, \dots, n$. Wir fassen die \mathbf{a}_i zu einer Matrix zusammen,

$$A = \begin{pmatrix} \mathbf{a}_1 \\ \vdots \\ \mathbf{a}_n \end{pmatrix} = \begin{pmatrix} a_{11} & \cdots & a_{1m} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nm} \end{pmatrix},$$

und setzen

$$A \cdot \vec{x} := \begin{pmatrix} \mathbf{a}_1 \bullet \mathbf{x} \\ \vdots \\ \mathbf{a}_n \bullet \mathbf{x} \end{pmatrix}.$$

So erhalten wir zwei Beschreibungen für die lineare Abbildung f :

$$f(\mathbf{x}) = (\mathbf{a}_1 \bullet \mathbf{x}, \dots, \mathbf{a}_n \bullet \mathbf{x}) \quad \text{und} \quad f(\vec{x}) = A \cdot \vec{x}.$$

Mit $\mathbf{z}_i(A) := \mathbf{a}_i$ bezeichnen wir die *Zeilen* von A , und mit

$$\vec{s}_j(A) := A \cdot \vec{e}_j = \begin{pmatrix} a_{1j} \\ \vdots \\ a_{nj} \end{pmatrix}$$

die *Spalten* von A .

Jede lineare Abbildung $f : \mathbb{R}^m \rightarrow \mathbb{R}^n$ definiert also eine Matrix $A = A(f) \in M_{n,m}(\mathbb{R})$, und umgekehrt liefert jede Matrix $A \in M_{n,m}(\mathbb{R})$ eine lineare Abbildung $f = f_A : \mathbb{R}^m \rightarrow \mathbb{R}^n$, durch $f_A(\vec{x}) = A \cdot \vec{x}$.

Beispiele.

1. Sei $f : \mathbb{R}^3 \rightarrow \mathbb{R}^2$ definiert durch $f(x_1, x_2, x_3) := (2x_1 + 3x_2 - x_3, 5x_2 + 7x_3)$. Das ist eine lineare Abbildung, die in der Form

$$f(\mathbf{x}) = (\mathbf{a}_1 \bullet \mathbf{x}, \mathbf{a}_2 \bullet \mathbf{x})$$

geschrieben werden kann, mit

$$\mathbf{a}_1 = (2, 3, -1) \quad \text{und} \quad \mathbf{a}_2 = (0, 5, 7).$$

Also ist

$$A = A(f) = \begin{pmatrix} 2 & 3 & -1 \\ 0 & 5 & 7 \end{pmatrix}.$$

Man bekommt diese Matrix auch über die Gleichung

$$A = (f(\vec{e}_1), f(\vec{e}_2), f(\vec{e}_3)).$$

Und zwar ist

$$f(1, 0, 0) = (2, 0), \quad f(0, 1, 0) = (3, 5) \quad \text{und} \quad f(0, 0, 1) = (-1, 7),$$

also

$$f(\vec{e}_1) = \begin{pmatrix} 2 \\ 0 \end{pmatrix}, \quad f(\vec{e}_2) = \begin{pmatrix} 3 \\ 5 \end{pmatrix} \quad \text{und} \quad f(\vec{e}_3) = \begin{pmatrix} -1 \\ 7 \end{pmatrix}.$$

Zusammengesetzt ergibt dies wieder die Matrix A .

2. Sei jetzt eine Matrix $A = \begin{pmatrix} 1 & 2 & 3 \\ -1 & 1 & 4 \\ 2 & 0 & 5 \end{pmatrix} \in M_{3,3}(\mathbb{R})$ gegeben. Dann ist

$$f_A \left(\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \right) = \begin{pmatrix} 1 & 2 & 3 \\ -1 & 1 & 4 \\ 2 & 0 & 5 \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} x_1 + 2x_2 + 3x_3 \\ -x_1 + x_2 + 4x_3 \\ 2x_1 + 5x_3 \end{pmatrix},$$

in Zeile-Schreibweise also

$$f_A(x_1, x_2, x_3) = (x_1 + 2x_2 + 3x_3, -x_1 + x_2 + 4x_3, 2x_1 + 5x_3).$$

Ist $f : V \rightarrow W$ eine lineare Abbildung zwischen zwei beliebigen Vektorräumen, so ist $f(\mathbf{0}) = f(\mathbf{0} + \mathbf{0}) = f(\mathbf{0}) + f(\mathbf{0})$, also $f(\mathbf{0}) = \mathbf{0}$. Das gilt immer, aber es kann darüber hinaus Vektoren $\mathbf{x} \neq \mathbf{0}$ mit $f(\mathbf{x}) = \mathbf{0}$ geben.

Definition:

Ist $f : V \rightarrow W$ linear, so setzt man

$$\text{Ker}(f) := \{\mathbf{x} \in V : f(\mathbf{x}) = \mathbf{0}\} \quad (\text{Kern von } f)$$

und

$$\text{Im}(f) := \{f(\mathbf{x}) : \mathbf{x} \in V\} \quad (\text{Bild von } f).$$

Satz

$\text{Ker}(f) \subset V$ und $\text{Im}(f) \subset W$ sind Unterräume.

BEWEIS: 1) Weil $f(\mathbf{0}) = \mathbf{0}$ ist, liegt $\mathbf{0}$ in $\text{Ker}(f)$ (und auch in $\text{Im}(f)$).

Sind $\mathbf{x}, \mathbf{y} \in \text{Ker}(f)$, so ist $f(\mathbf{x}) = \mathbf{0}$ und $f(\mathbf{y}) = \mathbf{0}$. Dann ist $f(\mathbf{x} + \mathbf{y}) = f(\mathbf{x}) + f(\mathbf{y}) = \mathbf{0} + \mathbf{0} = \mathbf{0}$, also $\mathbf{x} + \mathbf{y} \in \text{Ker}(f)$. Außerdem ist $f(\lambda\mathbf{x}) = \lambda f(\mathbf{x}) = \lambda\mathbf{0} = \mathbf{0}$, also auch $\lambda\mathbf{x} \in \text{Ker}(f)$.

2) Ist $\mathbf{v} = f(\mathbf{x}) \in \text{Im}(f)$ und $\mathbf{w} = f(\mathbf{y}) \in \text{Im}(f)$, so ist $\mathbf{v} + \mathbf{w} = f(\mathbf{x}) + f(\mathbf{y}) = f(\mathbf{x} + \mathbf{y}) \in \text{Im}(f)$ und $\lambda\mathbf{v} = \lambda f(\mathbf{x}) = f(\lambda\mathbf{x}) \in \text{Im}(f)$. ■

Satz

Sei $f : V \rightarrow W$ linear. f ist genau dann injektiv, wenn $\text{Ker}(f) = \{\mathbf{0}\}$ ist. f ist genau dann surjektiv, wenn $\text{Im}(f) = W$ ist.

BEWEIS: Die zweite Aussage ist trivial. Zur ersten Aussage:

Ist f injektiv und $\mathbf{x} \in \text{Ker}(f)$, so ist $f(\mathbf{x}) = \mathbf{0} = f(\mathbf{0})$, also $\mathbf{x} = \mathbf{0}$.

Sei umgekehrt $\text{Ker}(f) = \{\mathbf{0}\}$. Ist $f(\mathbf{x}_1) = f(\mathbf{x}_2)$, so ist $f(\mathbf{x}_1 - \mathbf{x}_2) = f(\mathbf{x}_1) - f(\mathbf{x}_2) = \mathbf{0}$, also $\mathbf{x}_1 - \mathbf{x}_2 = \mathbf{0}$ und damit $\mathbf{x}_1 = \mathbf{x}_2$. Das bedeutet, daß f injektiv ist. ■

Beispiel.

Sei $f : \mathbb{R}^3 \rightarrow \mathbb{R}^2$ definiert durch

$$f(x_1, x_2, x_3) := (x_1 - x_2, x_2 - x_3).$$

Dann ist $f(x_1, x_2, x_3) = (0, 0) \iff ((x_1 = x_2) \wedge (x_2 = x_3))$, also

$$\text{Ker}(f) = \{(x, x, x) : x \in \mathbb{R}\}.$$

Das ist ein 1-dimensionaler Unterraum (mit Basis $\{(1, 1, 1)\}$). Also ist f nicht injektiv.

Definieren wir dagegen $g : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ durch

$$g(x_1, x_2, x_3) := (x_1 - x_2, x_2 - x_3, x_1 + x_3),$$

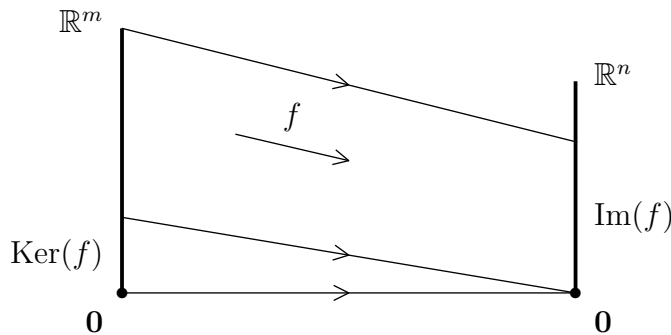
so ist $g(x_1, x_2, x_3) = (0, 0, 0) \iff ((x_1 = x_2 = x_3) \wedge (x_1 + x_3 = 0))$, also

$$\text{Ker}(g) = \{(x, x, x) : (x \in \mathbb{R}) \wedge (x + x = 0)\} = \{(0, 0, 0)\}.$$

Damit ist g injektiv.

Man beachte, daß der Begriff „Kern“ nur bei **linearen** Abbildungen benutzt werden darf! Ist $f : \mathbb{R} \rightarrow \mathbb{R}$ die Funktion $f(x) = x^2$, so ist zwar $f(x) = 0 \iff x = 0$, aber f ist trotzdem nicht injektiv.

Manchmal ist die folgende symbolische Darstellung einer linearen Abbildung hilfreich:



Dimensionsformel

Sei $f : V \rightarrow W$ eine lineare Abbildung zwischen endlich-dimensionalen Vektorräumen. Dann gilt:

$$\dim(\text{Ker}(f)) + \dim(\text{Im}(f)) = \dim(V).$$

BEWEIS: Es gibt eine Basis $\{\mathbf{a}_1, \dots, \mathbf{a}_k\}$ von $\text{Ker}(f)$, die wir zu einer Basis $\{\mathbf{a}_1, \dots, \mathbf{a}_k, \mathbf{a}_{k+1}, \dots, \mathbf{a}_m\}$ von V ergänzen können.

Ist jetzt $\mathbf{y} \in \text{Im}(f)$, so gibt es ein $\mathbf{x} = \alpha_1 \mathbf{a}_1 + \dots + \alpha_m \mathbf{a}_m \in V$ mit $f(\mathbf{x}) = \mathbf{y}$. Weil aber $f(\mathbf{a}_i) = \mathbf{0}$ für $i = 1, \dots, k$ ist, folgt:

$$\mathbf{y} = \alpha_{k+1} f(\mathbf{a}_{k+1}) + \dots + \alpha_m f(\mathbf{a}_m).$$

Das bedeutet, daß $E := \{f(\mathbf{a}_{k+1}), \dots, f(\mathbf{a}_m)\}$ ein Erzeugendensystem von $\text{Im}(f)$ ist. Wir zeigen, daß E sogar linear unabhängig, also eine Basis von $\text{Im}(f)$ ist.

Ist nämlich $\mathbf{0} = \sum_{i=k+1}^m \beta_i f(\mathbf{a}_i) = f\left(\sum_{i=k+1}^m \beta_i \mathbf{a}_i\right)$, so liegt $\sum_{i=k+1}^m \beta_i \mathbf{a}_i$ in $\text{Ker}(f)$, und es gibt reelle Zahlen γ_i , $i = 1, \dots, k$, so daß gilt:

$$\sum_{i=k+1}^m \beta_i \mathbf{a}_i = \sum_{i=1}^k \gamma_i \mathbf{a}_i,$$

also

$$\sum_{i=1}^k \gamma_i \mathbf{a}_i + \sum_{i=k+1}^m (-\beta_i) \mathbf{a}_i = \mathbf{0}.$$

Weil die \mathbf{a}_i linear unabhängig sind, müssen alle Koeffizienten verschwinden, insbesondere ist $\beta_{k+1} = \dots = \beta_m = 0$. Damit ist alles gezeigt. ■

Ist $f : \mathbb{R}^m \rightarrow \mathbb{R}^n$ eine lineare Abbildung und $\mathbf{b} \in \mathbb{R}^n$ ein fester Vektor, so bezeichnet man die Gleichung

$$\boxed{f(\mathbf{x}) = \mathbf{b}}$$

als *lineares Gleichungssystem*.

Ist $A = A(f) \in M_{n,m}(\mathbb{R})$ die Matrix zur linearen Abbildung f , so erhält die Gleichung die Form

$$\boxed{A \cdot \vec{x} = \vec{b}}$$

Ausführlich hingeschrieben ergibt das das Gleichungssystem

$$\boxed{\begin{array}{ccccccc} a_{11}x_1 & + & \cdots & + & a_{1m}x_m & = & b_1 \\ & & \vdots & & & & \vdots \\ a_{n1}x_1 & + & \cdots & + & a_{nm}x_m & = & b_n \end{array}}$$

Bemerkungen.

1. Die Gleichung ist genau dann lösbar, wenn $\mathbf{b} \in \text{Im}(f)$ ist.
2. Ist $\mathbf{x}_0 \in \mathbb{R}^m$ eine feste Lösung, so gilt für $\mathbf{x} \in \mathbb{R}^m$:

$$\begin{aligned} \mathbf{x} \text{ Lösung} &\iff f(\mathbf{x}) = f(\mathbf{x}_0) \\ &\iff f(\mathbf{x} - \mathbf{x}_0) = \mathbf{0} \\ &\iff \mathbf{x} - \mathbf{x}_0 \in \text{Ker}(f). \end{aligned}$$

Also ist die Lösungsmenge der affine Raum $\mathbf{x}_0 + \text{Ker}(f)$.

Die Gleichung $f(\mathbf{x}) = \mathbf{0}$ nennt man das *zugehörige homogene Gleichungssystem*. Hier ist die Lösungsmenge der Unterraum $\text{Ker}(f)$.

Jetzt verwenden wir die Matrizen-Schreibweise: $A \cdot \vec{x} = \vec{b}$.

Behauptung. $\text{Im}(f_A)$ wird von den Spalten $\vec{s}_1(A), \dots, \vec{s}_m(A)$ erzeugt.

BEWEIS: \vec{y} liegt genau dann in $\text{Im}(f_A)$, wenn es ein $\vec{x} \in \mathbb{R}^m$ mit $A \cdot \vec{x} = \vec{y}$ gibt. Schreibt man $\vec{x} = x_1 \vec{e}_1 + \dots + x_m \vec{e}_m$, so ist

$$\vec{y} = A \cdot \left(\sum_{i=1}^m x_i \vec{e}_i \right) = \sum_{i=1}^m x_i (A \cdot \vec{e}_i) = \sum_{i=1}^m x_i \vec{s}_i(A). \quad \blacksquare$$

Definition:

Die Zahl $\text{rg}(A) := \dim(\text{Im}(f_A))$ nennt man den *Rang* von A .

Es ist klar, daß $\text{rg}(A)$ die Maximalzahl linear unabhängiger Spalten von A ist.

Im Folgenden bezeichnen wir den (affinen) Lösungsraum der Gleichung $A \cdot \vec{x} = \vec{b}$ mit $\text{Lös}(A, \vec{b})$.

Lösungskriterium für lineare Systeme

$A \cdot \vec{x} = \vec{b}$ ist genau dann lösbar, wenn $\text{rg}(A, \vec{b}) = \text{rg}(A)$ ist. In diesem Fall ist $\text{Lös}(A, \vec{b})$ ein affiner Raum der Dimension $m - \text{rg}(A)$.

BEWEIS: Die Gleichung ist genau dann lösbar, wenn \vec{b} in $\text{Im}(f_A)$ liegt. Das ist genau dann der Fall, wenn \vec{b} Linearkombination von $\vec{s}_1(A), \dots, \vec{s}_m(A)$ ist. Dann ist der Rang der erweiterten Matrix (A, \vec{b}) nicht größer als der Rang der Matrix A selbst. Umgekehrt folgt aus $\text{rg}(A, \vec{b}) = \text{rg}(A)$, daß \vec{b} von den Spalten von A linear abhängt. ■

Folgendes Programm muß also ausgeführt werden:

1. Finde zunächst **eine** „partikuläre“ Lösung \vec{x}_0 der inhomogenen Gleichung $A \cdot \vec{x} = \vec{b}$.
2. Bestimme eine Basis $\{\vec{a}_1, \dots, \vec{a}_k\}$ von $\text{Ker}(f)$, wobei $k = m - \text{rg}(A)$ ist.

Dann ist $\text{Lös}(A, \vec{b}) = \{\vec{x} = \vec{x}_0 + t_1 \vec{a}_1 + \dots + t_k \vec{a}_k \mid t_i \in \mathbb{R} \text{ für } i = 1, \dots, k\}$.

Zur Durchführung dieses Programms benutzen wir das

Eliminationsverfahren von Gauß

Es geht darum, eine Variable nach der anderen aus dem Gleichungssystem zu eliminieren, bis nur noch eine Gleichung mit einer Unbekannten übrig bleibt, die dann problemlos zu lösen ist. Anschließend werden durch „Rückwärtseinsetzen“ sukzessive alle anderen Variablen bestimmt.

Sind $\mathbf{a}_i = \mathbf{z}_i(A)$ die Zeilen von A , so kann man das Gleichungssystem auch folgendermaßen schreiben:

$$\begin{aligned} \mathbf{a}_1 \bullet \mathbf{x} &= b_1 \\ &\vdots \\ \mathbf{a}_n \bullet \mathbf{x} &= b_n \end{aligned}$$

Die folgenden „**Zeilenoperationen**“ lassen die Lösungsmenge invariant, sofern sie auf die erweiterte Matrix (A, \vec{b}) angewandt werden:

(I) **Multiplikation der i -ten Zeile mit einer Zahl $\lambda \neq 0$.**

Denn die Gleichung $\mathbf{a}_i \bullet \mathbf{x} = b_i$ besitzt die gleiche Lösungsmenge wie die Gleichung $(\lambda \mathbf{a}_i) \bullet \mathbf{x} = \lambda b_i$.

(II) **Addition der k -ten Zeile zur i -ten Zeile.**

Denn das Gleichungssystem

$$\begin{aligned} \mathbf{a}_i \bullet \mathbf{x} &= b_i \\ \mathbf{a}_k \bullet \mathbf{x} &= b_k \end{aligned}$$

besitzt die gleiche Lösungsmenge wie das Gleichungssystem

$$\begin{aligned} (\mathbf{a}_i + \mathbf{a}_k) \bullet \mathbf{x} &= b_i + b_k \\ \mathbf{a}_k \bullet \mathbf{x} &= b_k \end{aligned}$$

Beliebige Kombinationen der Operationen (I) und (II) sind erlaubt. Da die Multiplikation mit -1 möglich ist, erhalten wir auch Subtraktionen von Zeilen. Und sogar Vertauschungen von Zeilen sind möglich:

$$\begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} x \\ -y \end{pmatrix} \mapsto \begin{pmatrix} x \\ x - y \end{pmatrix} \mapsto \begin{pmatrix} x - (x - y) \\ x - y \end{pmatrix} \mapsto \begin{pmatrix} y \\ x - y \end{pmatrix} \mapsto \begin{pmatrix} y \\ x \end{pmatrix}$$

Als „**Spaltenoperationen**“ lassen wir nur Vertauschungen von Spalten zu. Dazu müssen wir etwas weiter ausholen.

Unter einer *Permutation* (von $\{1, \dots, n\}$) versteht man eine bijektive Abbildung $\sigma : \{1, \dots, n\} \rightarrow \{1, \dots, n\}$. Wir schreiben σ auch in der Form

$$\sigma = \begin{pmatrix} 1 & 2 & \cdots & n \\ \sigma(1) & \sigma(2) & & \sigma(n) \end{pmatrix}.$$

Zum Beispiel steht $\begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix}$ für die Abbildung $\sigma : \{1, 2, 3\} \rightarrow \{1, 2, 3\}$ mit

$$\sigma(1) = 3, \quad \sigma(2) = 1 \quad \text{und} \quad \sigma(3) = 2.$$

Die Menge aller Permutationen bezeichnen wir mit S_n . Eine Permutation $\sigma \in S_n$ heißt *Vertauschung*, wenn es $i \neq j$ gibt, so daß $\sigma(i) = j$, $\sigma(j) = i$ und $\sigma(k) =$

k in allen anderen Fällen ist. Man kann zeigen, daß sich jede Permutation aus Vertauschungen zusammensetzt.

Jetzt können wir die dritte Sorte elementarer Umformungen einführen:

(III) Vertauschung von Spalten.

Ist $\sigma \in S_n$ eine Permutation, so wird $P_\sigma : \mathbb{R}^n \rightarrow \mathbb{R}^n$ definiert durch

$$P_\sigma(x_1, \dots, x_n) := (x_{\sigma(1)}, \dots, x_{\sigma(n)}).$$

Dann ist

$$\begin{aligned} P_\sigma(\mathbf{a}_i) \bullet P_\sigma(\mathbf{x}) &= a_{i\sigma(1)}x_{\sigma(1)} + \dots + a_{i\sigma(n)}x_{\sigma(n)} \\ &= a_{i1}x_1 + \dots + a_{in}x_n \\ &= \mathbf{a}_i \bullet \mathbf{x}. \end{aligned}$$

Das bedeutet, daß eine Vertauschung von Spalten von A die Vertauschung der entsprechenden Komponenten der Lösungsvektoren bedingt.

Ist also z.B. der Vektor $\mathbf{y} = (y_1, y_2, y_3)$ eine Lösung der Gleichung

$$(\vec{s}_1(A), \vec{s}_2(A), \vec{s}_3(A)) \cdot \vec{x} = \vec{b},$$

so ist (y_3, y_2, y_1) eine Lösung der Gleichung

$$(\vec{s}_3(A), \vec{s}_2(A), \vec{s}_1(A)) \cdot \vec{x} = \vec{b}.$$

Rang-Erhaltungssatz

Bei elementaren Umformungen vom Typ (I), (II) oder (III) ändert sich der Rang einer Matrix nicht.

BEWEIS: Der Rang von A ist die Maximalzahl linear unabhängiger Spalten von A . Für Spaltenvertauschungen ist deshalb die Aussage des Satzes trivial.

Bei Zeilenoperationen bleibt der Lösungsraum $\text{Ker}(f_A)$ der homogenen Gleichung $A \cdot \vec{x} = \vec{0}$ erhalten, also auch $\text{rg}(A) = m - \dim(\text{Ker}(f_A))$. ■

Definition:

Eine Matrix $A \in M_{n,m}(\mathbb{R})$ soll *r-spezial* genannt werden, wenn sie folgendermaßen in Kästchen aufgeteilt werden kann:

$$A = \left(\begin{array}{c|c} D & B \\ \hline 0 & C \end{array} \right), \text{ mit } B \in M_{r,m-r}(\mathbb{R}), C \in M_{n-r,m-r}(\mathbb{R}) \quad \text{und}$$

$$D = \begin{pmatrix} a_{11} & \cdots & a_{1r} \\ \vdots & \ddots & \vdots \\ 0 & \cdots & a_{rr} \end{pmatrix} \in M_{r,r}(\mathbb{R}), a_{11}, \dots, a_{rr} \neq 0.$$

Bemerkungen.

1. Der Begriff „*r-spezial*“ ist kein Standardbegriff. Ich verwende ihn hier nur, um das Gauß-Verfahren besser erklären zu können.
2. Jede Matrix ist 0-spezial. Dann sind D und B nicht vorhanden, und es ist $A = C$.
3. Ist $r = n = m$ und A *r-spezial*, so ist A eine „obere Dreiecksmatrix“.
4. Es ist stets $r \leq \min(n, m)$.
5. Ist A *r-spezial* und zusätzlich $C = 0$ (oder – im Falle $r = n$ – nicht vorhanden), so nennen wir A eine *Gauß-Matrix*. Auch dieser Begriff ist kein Standard-Begriff. Eine (*r-speziale*) Gauß-Matrix hat stets den Rang r .

Ist $A \in M_{n,m}(\mathbb{R})$ eine (*r-speziale*) Gauß-Matrix, so kann man das Gleichungssystem $A \cdot \vec{x} = \vec{b}$ genau dann lösen, wenn $b_{r+1} = \dots = b_n = 0$ ist.

Die Bedingung ist notwendig, denn auf der linken Seite stehen in der $(r+1)$ -ten bis zur n -ten Zeile nur Nullen. Sie ist aber auch hinreichend, denn das Gleichungssystem reduziert sich nun auf

$$\begin{array}{cccccccc} a_{11}x_1 & + & a_{12}x_2 & + & \cdots & + & a_{1r}x_r & + & a_{1,r+1}x_{r+1} & + & \cdots & + & a_{1m}x_m & = & b_1, \\ & & a_{22}x_2 & + & \cdots & + & a_{2r}x_r & + & a_{2,r+1}x_{r+1} & + & \cdots & + & a_{2m}x_m & = & b_2, \\ & & & & \ddots & & & & & & & & & & \vdots \\ & & & & & & a_{rr}x_r & + & a_{r,r+1}x_{r+1} & + & \cdots & + & a_{rm}x_m & = & b_r. \end{array}$$

Jetzt kann man für x_{r+1}, \dots, x_m beliebige Werte wählen und $y_i := b_i - a_{i,r+1}x_{r+1} - \dots - a_{im}x_m$ setzen. Dann erhält man das folgende Gleichungssystem:

$$\begin{array}{rcl}
 a_{11}x_1 + a_{12}x_2 + \cdots + a_{1r}x_r & = & y_1, \\
 a_{22}x_2 + \cdots + a_{2r}x_r & = & y_2, \\
 & \vdots & \\
 a_{rr}x_r & = & y_r.
 \end{array}$$

Dieses System kann man ganz einfach durch „Rückwärtseinsetzen“ lösen:

$$\begin{array}{rcl}
 x_r & = & \frac{1}{a_{rr}} \cdot y_r, \\
 x_{r-1} & = & \frac{1}{a_{1-r,1-r}} \cdot (y_{r-1} - a_{r-1,r}x_r), \\
 & \vdots & \\
 x_1 & = & \frac{1}{a_{11}} \cdot (y_1 - a_{12}x_2 - \cdots - a_{1r}x_r).
 \end{array}$$

Der folgende Satz erlaubt es nun, beliebige Gleichungssysteme zu lösen.

Gaußsches Eliminationsverfahren

Ist $A \in M_{n,m}(\mathbb{R})$ und $\text{rg}(A) = r$, so kann man A durch elementare Umformungen vom Typ (I), (II) und (III) in eine r -spezielle Gauß-Matrix umformen.

BEWEIS: A ist auf jeden Fall 0-speziell. Sei nun A schon k -speziell, für ein k mit $0 \leq k < r$, also

$$A = \left(\begin{array}{c|c} D_k & B_k \\ \hline 0 & C_k \end{array} \right), \text{ mit einer oberen Dreiecksmatrix } D_k \in M_{k,k}(\mathbb{R}).$$

Ist $C_k = 0$, so ist $\text{rg}(A) = k < r$. Das ist ein Widerspruch. Also muß es ein a_{ij} in C_k geben, das $\neq 0$ ist. Wir nennen es das *Pivot-Element*. Natürlich ist es nicht eindeutig bestimmt.

Durch Vertauschen von Zeilen und Spalten kann man erreichen, daß $a_{k+1,k+1} \neq 0$ ist.

Durch Subtraktion geeigneter Vielfache der $(k+1)$ -ten Zeile von den folgenden Zeilen kann man erreichen, daß $a_{i,k+1} = 0$ wird, für $i = k+2, \dots, n$. Das Ergebnis ist eine $(k+1)$ -spezielle Matrix.

Nach endlich vielen Schritten ist A r -speziell. Da noch immer $\text{rg}(A) = r$ ist, müssen die hinteren Spalten von den ersten r Spalten linear abhängen. Das bedeutet, daß $C_r = 0$ (oder nicht mehr vorhanden) ist. A ist zu einer Gauß-Matrix geworden. ■

Da sich der Rang einer Matrix bei den elementaren Umformungen nicht ändert, braucht man ihn nicht vorher zu kennen, er ergibt sich automatisch aus dem Eliminationsverfahren.

Beispiel.

Die Matrix $\begin{pmatrix} 1 & 1 & 1 \\ 1 & -1 & -1 \\ 1 & 3 & 3 \end{pmatrix}$ ist 0-speziell. Damit ist hier $A = C_0$, die Matrizen D_0, B_0 sind nicht vorhanden.

Da $a_{11} = 1 \neq 0$ ist, subtrahieren wir die erste Zeile von der 2. und der 3. Zeile. Das ergibt die neue Matrix

$$A_1 = \begin{pmatrix} 1 & 1 & 1 \\ 0 & -2 & -2 \\ 0 & 2 & 2 \end{pmatrix}.$$

Jetzt ist $D_1 = (1)$, $B_1 = (1, 1)$ und $C_1 = \begin{pmatrix} -2 & -2 \\ 2 & 2 \end{pmatrix}$.

Da jetzt $a_{22} = -2 \neq 0$ ist, addieren wir die 2. Zeile zu der 3. Zeile. Dann erhalten wir:

$$A_2 = \begin{pmatrix} 1 & 1 & 1 \\ 0 & -2 & -2 \\ 0 & 0 & 0 \end{pmatrix},$$

also $D_2 = \begin{pmatrix} 1 & 1 \\ 0 & -2 \end{pmatrix}$, $B_2 = \begin{pmatrix} 1 \\ -2 \end{pmatrix}$ und $C_2 = (0)$. Damit ist das Verfahren abgeschlossen, A_2 ist eine 2-spezielle Gauß-Matrix. Insbesondere ist $\text{rg}(A) = 2$ und $\dim(\text{Ker}(f_A)) = 3 - 2 = 1$.

Jetzt soll folgendes Gleichungssystem gelöst werden:

$$\begin{pmatrix} 1 & 1 & 1 \\ 1 & -1 & -1 \\ 1 & 3 & 3 \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 4 \\ 0 \\ 8 \end{pmatrix}.$$

Wenden wir die obigen elementaren Umformungen auf die erweiterte Matrix an, so erhalten wir das Gleichungssystem

$$\begin{pmatrix} 1 & 1 & 1 \\ 0 & -2 & -2 \\ 0 & 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 4 \\ -4 \\ 0 \end{pmatrix}.$$

Um eine spezielle Lösung des inhomogenen Systems zu erhalten, können wir für x_3 irgend einen Wert einsetzen. Der Einfachheit halber nehmen wir $x_3 = 0$. Dann bleibt nur noch folgendes Gleichungssystem übrig:

$$\begin{pmatrix} 1 & 1 \\ 0 & -2 \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 4 \\ -4 \end{pmatrix}.$$

Rückwärtseinsetzen liefert $x_2 = 2$ und $x_1 = 4 - 2 = 2$, also die Lösung $\mathbf{x}_0 = (2, 2, 0)$. Jetzt müssen wir eine Basis des Lösungsraumes der zugehörigen homogenen Gleichung (also von $\text{Ker}(f_A)$) finden. Dazu müssen wir nur **einen** Lösungsvektor $\mathbf{a} \neq \mathbf{0}$ des folgenden Gleichungssystems finden:

$$\begin{pmatrix} 1 & 1 & 1 \\ 0 & -2 & -2 \\ 0 & 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

Wieder können wir x_3 beliebig wählen, und um sicherzustellen, daß nicht der Nullvektor herauskommt, setzen wir $x_3 = 1$. Dann ist

$$x_1 + x_2 = -1 \quad \text{und} \quad -2x_2 = 2,$$

also $x_2 = -1$ und $x_1 = 0$. Das ergibt $\mathbf{a} = (0, -1, 1)$.

Als Lösungsgesamtheit erhalten wir deshalb

$$\{(x_1, x_2, x_3) = (2, 2, 0) + t(0, -1, 1) : t \in \mathbb{R}\} = \{(2, 2 - t, t) : t \in \mathbb{R}\}.$$

Das im Beispiel vorgeführte Verfahren soll noch allgemein formuliert werden. Wir gehen von folgender Gleichung aus:

$$(D_r | B_r) \cdot \begin{pmatrix} \vec{x}^* \\ \vec{x}^{**} \end{pmatrix} = \vec{b},$$

mit $\mathbf{x} = (x_1, \dots, x_r)$, $\mathbf{x}^{**} = (x_{r+1}, \dots, x_m)$, $\vec{b} \in \mathbb{R}^r$ und $D_r = \begin{pmatrix} a_{11} & \cdots & a_{1r} \\ \vdots & \ddots & \vdots \\ 0 & \cdots & a_{rr} \end{pmatrix}$.

Man kann die Gleichung wie folgt umformen:

$$D_r \cdot \vec{x}^* = \vec{b} - B_r \cdot \vec{x}^{**}.$$

Um eine spezielle Lösung der inhomogenen Gleichung zu erhalten, setzen wir $\vec{x}^{**} = \vec{0}$. Dann erhält man \vec{x}^* aus $D_r \cdot \vec{x}^* = \vec{b}$ durch Rückwärtseinsetzen.

Um eine Basis des Lösungsraumes der homogenen Gleichung zu finden, setzen wir für \vec{x}^{**} nacheinander die Einheitsvektoren $\vec{e}_1, \dots, \vec{e}_{m-r}$ ein und berechnen die zugehörigen Vektoren $\vec{x}_1^*, \dots, \vec{x}_{m-r}^*$ aus $D_r \cdot \vec{x}_i^* = \vec{b} - B_r \cdot \vec{e}_i$. Das stellt sicher, daß die Lösungsvektoren $\mathbf{a}_1 = (\mathbf{x}_1^*, \mathbf{e}_1), \dots, \mathbf{a}_{m-r} = (\mathbf{x}_{m-r}^*, \mathbf{e}_{m-r})$ linear unabhängig sind.

Beispiel.

Wir wollen jetzt ein etwas größeres Gleichungssystem betrachten:

$$\begin{aligned}
 x_1 + 3x_2 - 4x_3 + 3x_4 &= 9, \\
 3x_1 + 9x_2 - 2x_3 - 11x_4 &= -3, \\
 4x_1 + 12x_2 - 6x_3 - 8x_4 &= 6, \\
 2x_1 + 6x_2 + 2x_3 - 14x_4 &= -12.
 \end{aligned}$$

Für die systematische Anwendung des Gauß'schen Eliminationsverfahrens benutzen wir folgendes Schema:

x_1	x_2	x_3	x_4	b_i	
1	3	-4	3	9	
3	9	-2	-11	-3	
4	12	-6	-8	6	
2	6	2	-14	-12	
1	3	-4	3	9	bleibt stehen
0	0	10	-20	-30	(-3 × 1. Zeile)
0	0	10	-20	-30	(-4 × 1. Zeile)
0	0	10	-20	-30	(-2 × 1. Zeile)

Die jetzt erhaltene Matrix ist 1-speziell, und ein Pivot-Element findet sich in der 3. Spalte. Also muß eine Spalten-Vertauschung vorgenommen werden.

x_1	x_2	x_3	x_4	b_i	
1	3	-4	3	9	
0	0	10	-20	-30	
0	0	10	-20	-30	
0	0	10	-20	-30	
x_1	x_3	x_2	x_4		Spaltenvertauschung
1	-4	3	3	9	
0	10	0	-20	-30	
0	10	0	-20	-30	
0	10	0	-20	-30	
1	-4	3	3	9	
0	10	0	-20	-30	bleibt stehen
0	0	0	0	0	(-2. Zeile)
0	0	0	0	0	(-2. Zeile)

Da jetzt links eine Gauß-Matrix steht, kann das System aufgelöst werden.

Für die spezielle Lösung setzen wir $x_2 = x_4 = 0$ und erhalten die Gleichungen

$$x_1 - 4x_3 = 9 \quad \text{und} \quad 10x_3 = -30,$$

also $x_3 = -3$ und $x_1 = -3$.

Eine Basis für den Lösungsraum des zugehörigen homogenen Systems erhält man, indem man für (x_2, x_4) die beiden möglichen Einheitsvektoren $(1, 0)$ und $(0, 1)$ einsetzt und dann die zugehörigen Werte von x_1 und x_3 bestimmt:

$x_2 = 1$ und $x_4 = 0$ ergibt

$$\begin{aligned}x_1 - 4x_3 + 3 &= 0 \\ \text{und } 10x_3 &= 0,\end{aligned}$$

also $x_1 = -3$ und $x_3 = 0$.

$x_2 = 0$ und $x_4 = 1$ ergibt

$$\begin{aligned}x_1 - 4x_3 + 3 &= 0 \\ \text{und } 10x_3 - 20 &= 0,\end{aligned}$$

also $x_1 = 5$ und $x_3 = 2$.

Die allgemeine Lösung des inhomogenen Systems ist also gegeben durch

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} -3 \\ 0 \\ -3 \\ 0 \end{pmatrix} + \alpha \cdot \begin{pmatrix} -3 \\ 1 \\ 0 \\ 0 \end{pmatrix} + \beta \cdot \begin{pmatrix} 5 \\ 0 \\ 2 \\ 1 \end{pmatrix} = \begin{pmatrix} -3 - 3\alpha + 5\beta \\ \alpha \\ -3 + 2\beta \\ \beta \end{pmatrix},$$

mit $\alpha, \beta \in \mathbb{R}$.

§ 4 Komplexe Zahlen

Inhalt:

Der Körper der komplexen Zahlen, Real- und Imaginärteil, konjugiert-komplexe Zahl, Betrag und Inverses.

Komplexwertige Funktionen einer reellen Veränderlichen, Argument und Polardarstellung, Eulersche Formel, Moivresche Formel, komplexe Wurzeln.

Der Fundamentalsatz der Algebra (Zitat), komplexe Nullstellen reeller Polynome.

Die Gleichung $x^2 + 1 = 0$ hat in \mathbb{R} keine Lösung. Nun hat aber z.B. auch die Gleichung $x^2 - 2 = 0$ in \mathbb{Q} keine Lösung, und dieses Problem kann dadurch beseitigt werden, daß man zu dem größeren Zahlenbereich \mathbb{R} übergeht.

Also stellt sich die Frage: Gibt es einen größeren Zahlenbereich als \mathbb{R} , in dem vielleicht die Gleichung $x^2 + 1 = 0$ lösbar ist? Oder genauer: Gibt es einen Körper K mit folgenden Eigenschaften:

1. $\mathbb{R} \subset K$.
2. Die Addition und die Multiplikation können mit all ihren Regeln auf K fortgesetzt werden.
3. Es gibt ein Element $j \in K$ mit $j^2 = -1$.

Auf den ersten Blick erscheint das ziemlich unsinnig. Aber nimmt man die Existenz von K trotzdem einmal an, so erhält man (für $a, b, c, d \in \mathbb{R}$) folgende Regeln:

$$\begin{aligned} (a + bj) + (c + dj) &= (a + c) + (b + d)j, \\ \text{und } (a + bj) \cdot (c + dj) &= (ac - bd) + (ad + bc)j. \end{aligned}$$

Was sagt uns das? Wenn es den gewünschten Körper K gibt, und wenn jedes Element von K in der Form $a + bj$ mit $a, b \in \mathbb{R}$ dargestellt werden kann, so wissen wir schon, wie wir zwei Elemente von K addieren und multiplizieren müssen. Aber dann sieht es so aus, als müßten wir die Elemente von K in einem 2-dimensionalen Vektorraum suchen, also in einer Ebene. Dafür gibt es auch eine plausible anschauliche Erklärung: Wenn die Multiplikation mit -1 (in \mathbb{R}) eine Drehung um 180° bedeutet und wenn $j^2 = -1$ ist, so sollte die Multiplikation mit j eine Drehung um 90° bedeuten. Das zwingt uns aus der Zahlengerade hinaus in die Ebene.

Jetzt haben wir genug Motivation für die folgende

Definition:

Unter der Menge der *komplexen Zahlen* versteht man die Ebene \mathbb{R}^2 , auf der neben der vektoriellen Addition zusätzlich eine Multiplikation gegeben ist:

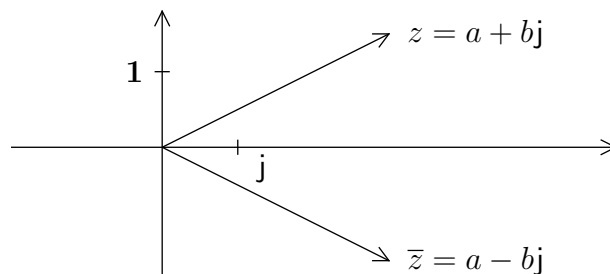
$$(a, b) \cdot (c, d) := (ac - bd, ad + bc).$$

Das Element $(1, 0)$ wird mit $\mathbf{1}$ bezeichnet, das Element $(0, 1)$ mit \mathbf{j} .

Die Menge der komplexen Zahlen wird auch mit dem Symbol \mathbb{C} bezeichnet. Jedes Element $z \in \mathbb{C}$ kann auf eindeutige Weise in der Form $z = a + b\mathbf{j}$ geschrieben werden, mit $a, b \in \mathbb{R}$. Man nennt dann $\operatorname{Re}(z) := a$ den *Realteil* und $\operatorname{Im}(z) := b$ den *Imaginärteil* von z . Das Element \mathbf{j} wird auch als *imaginäre Einheit* bezeichnet.

Die reellen Zahlen bilden eine Teilmenge der komplexen Zahlen (nämlich genau diejenigen komplexen Zahlen, deren Imaginärteil = 0 ist). Da wir die Regeln der Vektorrechnung schon kennen, brauchen wir uns nur noch mit der multiplikativen Struktur von \mathbb{C} zu beschäftigen. Die 1 ist sicher ein neutrales Element für die Multiplikation. Assoziativität und Kommutativität der Multiplikation kann man – mit genügend Geduld – nachrechnen, genauso das Distributivgesetz. Nun müssen wir noch das Reziproke zu einer komplexen Zahl $z \neq 0$ finden. Eine direkte Herleitung ist etwas mühsam, deshalb verwenden wir einen einfachen Trick:

Ist $z = a + b\mathbf{j} \in \mathbb{C}$, so nennt man $\bar{z} := a - b\mathbf{j}$ die zu z *konjugierte komplexe Zahl*.



Es gilt:

$$1. \quad \overline{z + w} = \bar{z} + \bar{w}.$$

$$\begin{aligned} \overline{(a + b\mathbf{j}) + (c + d\mathbf{j})} &= \overline{(a + c) + (b + d)\mathbf{j}} = (a + c) - (b + d)\mathbf{j} \\ &= (a - b\mathbf{j}) + (c - d\mathbf{j}) = \overline{a + b\mathbf{j}} + \overline{c + d\mathbf{j}}. \end{aligned}$$

$$2. \quad \overline{z \cdot w} = \bar{z} \cdot \bar{w}.$$

$$\begin{aligned} \overline{(a + b\mathbf{j}) \cdot (c + d\mathbf{j})} &= \overline{(ac - bd) + (ad + bc)\mathbf{j}} = (ac - bd) - (ad + bc)\mathbf{j} \\ \text{und} \\ \overline{(a + b\mathbf{j})} \cdot \overline{(c + d\mathbf{j})} &= (a - b\mathbf{j}) \cdot (c - d\mathbf{j}) = (ac - bd) - (ad + bc)\mathbf{j}. \end{aligned}$$

3. $\overline{\overline{z}} = z$.

4. Ist $z = a + bj$, so ist $z \cdot \overline{z} = (a + bj) \cdot (a - bj) = a^2 + b^2 \geq 0$.
Ist $z \neq 0$, so ist sogar $z \cdot \overline{z} > 0$.

5. Realteil und Imaginärteil einer komplexen Zahl sind gegeben durch

$$\operatorname{Re}(z) = \frac{1}{2}(z + \overline{z}) \text{ und } \operatorname{Im}(z) = \frac{1}{2j}(z - \overline{z}).$$

Die reelle Zahl $|z| := +\sqrt{z\overline{z}} \geq 0$ nennt man den *Betrag* der komplexen Zahl z . Ist $z = a + bj$, so ist

$$|z| = \sqrt{a^2 + b^2} = \|(a, b)\|$$

die euklidische Norm und besitzt deshalb auch die gleichen Eigenschaften.

Ist nun $z \neq 0$, so ist $z\overline{z} = |z|^2 > 0$, und es gilt:

$$1 = \frac{z\overline{z}}{z\overline{z}} = z \cdot \frac{\overline{z}}{|z|^2}.$$

Also ist

$$z^{-1} = \frac{\overline{z}}{|z|^2}, \text{ für } z \neq 0.$$

Damit ist \mathbb{C} ein Körper. Allerdings kann \mathbb{C} nicht wie \mathbb{R} angeordnet werden, denn wenn es eine Anordnung auf \mathbb{C} gäbe, dann müsste $1 \cdot 1 = 1$ und $j \cdot j = -1$ positiv sein, also auch $0 = 1 + (-1)$. Aber das widerspricht den Regeln einer Anordnung.

Definition:

Sei $I \subset \mathbb{R}$ ein Intervall. Eine Abbildung $f : I \rightarrow \mathbb{C}$ wird auch als *komplexwertige Funktion* bezeichnet.

Sie kann immer in der Form $f = g + jh$ in Realteil und Imaginärteil zerlegt werden. g und h sind dabei gewöhnliche reelle Funktionen. Man nennt dann f z.B. stetig, wenn g und h es sind.

Ein wichtiges Beispiel einer solchen komplexwertigen Funktion wollen wir sogleich einführen:

Ist $z = a + bj$ eine komplexe Zahl mit $|z| = 1$, so ist $a^2 + b^2 = 1$ und es gibt eine eindeutig bestimmte Zahl $t \in [0, 2\pi)$ mit

$$\cos(t) = a \text{ und } \sin(t) = b.$$

Für beliebige komplexe Zahlen $z \neq 0$ erhält man dementsprechend eine Darstellung

$$z = |z| \cdot (\cos t + j \sin t).$$

Das ist die sogenannte „Polar-Darstellung“ von z . Die Zahl $\arg(z) := t$ nennt man das *Argument* von z . Erlaubt man, daß t beliebig in \mathbb{R} gewählt wird, so ist das Argument nur bis auf 2π eindeutig bestimmt. Für $z = 0$ kann man überhaupt keinen Winkel ermitteln, deswegen haben wir diesen Fall ausgeschlossen.

Nun folgt eine auf den ersten Blick befremdliche Definition:

Definition:

Für $t \in \mathbb{R}$ sei $e^{jt} := \cos(t) + j \sin(t)$.

Die Rechtfertigung liefert der folgende Satz:

Rechenregeln für e^{jt}

1. $e^{j \cdot 0} = 1$.

2. $e^{j(t+s)} = e^{jt} \cdot e^{js}$.

BEWEIS: 1) $\cos(0) = 1$ und $\sin(0) = 0$.

2) Mit dem Additionstheorem erhält man:

$$\begin{aligned} \cos(t+s) + j \sin(t+s) &= \\ &= \cos(t)\cos(s) - \sin(t)\sin(s) + j[\sin(t)\cos(s) + \cos(t)\sin(s)] \\ &= [\cos(t) + j \sin(t)] \cdot [\cos(s) + j \sin(s)]. \end{aligned}$$

■

Bemerkung. Die Beziehung

$$e^{jt} = \cos(t) + j \sin(t)$$

wird als *Eulersche Formel* bezeichnet.

Das Rechnen mit Winkelfunktionen wird auf diese Weise stark vereinfacht. Z.B. gilt:

Moivre'sche Formel

$$(\cos t + j \sin t)^n = \cos(nt) + j \sin(nt).$$

BEWEIS:

$$(e^{jt})^n = e^{jnt}.$$

■

Interessant ist auch die folgende „Weltformel“:

$$e^{j\pi} + 1 = 0.$$

Sie verbindet die wichtigsten Konstanten der Mathematik miteinander: $0, 1, j, \pi$ und e . Der Beweis ist ganz simpel: Es ist $\cos(\pi) = -1$ und $\sin(\pi) = 0$. Die Bezeichnung „Weltformel“ ist also nur ein Scherz.

Nun versuchen wir, Wurzeln aus komplexen Zahlen zu ziehen.

Existenz der n -ten Wurzeln aus 1

Die Gleichung $z^n = 1$ hat in \mathbb{C} genau n Lösungen, nämlich

$$\zeta_{k,n} := e^{\frac{k}{n} \cdot 2\pi j}, \quad k = 0, 1, \dots, n-1.$$

BEWEIS: Die Punkte $\zeta_{k,n} = \cos(k \cdot \frac{2\pi}{n}) + j \sin(k \cdot \frac{2\pi}{n})$, $k = 0, 1, \dots, n-1$, liegen auf den Ecken eines (dem Einheitskreis einbeschriebenen) regelmäßigen n -Ecks. Insbesondere sind sie alle verschieden.

Da $e^{k \cdot 2\pi j} = \cos(k \cdot 2\pi) + j \sin(k \cdot 2\pi) = 1$ ist, ist

$$(\zeta_{k,n})^n = e^{k \cdot 2\pi j} = 1 \text{ für } k = 0, 1, \dots, n-1.$$

Sei umgekehrt $w \in \mathbb{C}$ irgendeine Lösung der Gleichung $z^n = 1$. Dann ist $|w|^n = |w^n| = 1$, also $|w| = 1$, also $w = e^{jt}$ für ein geeignetes $t \in [0, 2\pi)$. Da außerdem $e^{jtn} = w^n = 1$ ist, muß gelten:

$$\cos(tn) = 1 \text{ und } \sin(tn) = 0.$$

Das ist nur möglich, wenn $tn \in \{2\pi k \mid k \in \mathbb{Z}\}$ ist. Da $t \in [0, 2\pi)$ liegt, kommen für tn nur die Werte $0, 2\pi, 4\pi, \dots, (n-1)2\pi$ in Frage. Also muß t von der Form $t = \frac{k}{n} \cdot 2\pi$ sein. ■

Definition:

Die Zahlen $\zeta_{k,n} := e^{\frac{k}{n} \cdot 2\pi j}$, $k = 0, 1, \dots, n-1$, heißen die n -ten Einheitswurzeln.

Man braucht übrigens zu jedem n jeweils nur die erste Einheitswurzel zu kennen, denn es ist ja $\zeta_{k,n} = (\zeta_{1,n})^k$.

Beispiel.

Sei $n = 3$. Es ist $\frac{2\pi}{3} = \text{Arcus}(120^\circ)$,

$$\cos(120^\circ) = -\frac{1}{2} \quad \text{und} \quad \sin(120^\circ) = \frac{1}{2}\sqrt{3}.$$

Daraus folgt:

$$\zeta_{0,3} = 1, \quad \zeta_{1,3} = -\frac{1}{2} + \frac{j}{2}\sqrt{3} \quad \text{und} \quad \zeta_{2,3} = (\zeta_{1,3})^2 = -\frac{1}{2} - \frac{j}{2}\sqrt{3}.$$

Satz

In \mathbb{C} besitzt jede Zahl $z \neq 0$ genau n n -te Wurzeln.

BEWEIS: Sei $z = re^{jt}$, mit $r = |z|$ und einem geeigneten $t \in [0, 2\pi)$. Dann setzen wir

$$z_k := \sqrt[n]{r} \cdot e^{\frac{jt}{n}} \cdot \zeta_{k,n}, \quad k = 0, 1, \dots, n-1.$$

Offensichtlich sind dies n verschiedene komplexe Zahlen z_k mit $z_k^n = z$.

Ist andererseits w irgendeine Lösung der Gleichung $w^n = z$, so ist $w^n = z_0^n$, also

$$(wz_0^{-1})^n = 1.$$

Das bedeutet, daß es eine n -te Einheitswurzel $\zeta_{k,n}$ gibt, so daß $w = z_0 \cdot \zeta_{k,n}$ ist. ■

Man kann also in \mathbb{C} nie von *der* n -ten Wurzel einer Zahl z sprechen, es gibt stets n verschiedene. Das gilt auch im Falle $n = 2$! Das Symbol \sqrt{z} ist also zweideutig. In \mathbb{R} haben wir dagegen die positive Lösung der Gleichung $x^2 = a$ als *die* Wurzel aus a definiert.

Der obige Satz enthält eine wichtige Aussage über komplexe Polynome.¹

¹Ein komplexes Polynom ist eine Zuordnung $z \mapsto p(z) = a_n z^n + a_{n-1} z^{n-1} + \dots + a_1 z + a_0$, also eine Funktion, die *komplexe Argumente* zuläßt. Genauer werden wir solche Funktionen erst im 3. Semester behandeln.

Das Polynom $z^n - w$ besitzt (für jedes $w \in \mathbb{C}$) eine Nullstelle.

Im Reellen ist das falsch, z.B. hat $x^2 + 1$ keine Nullstelle in \mathbb{R} . Das zeigt die Bedeutung des folgenden Satzes:

Fundamentalsatz der Algebra

Jedes nicht konstante komplexe Polynom hat in \mathbb{C} wenigstens eine Nullstelle.

Wir verzichten hier auf den nicht ganz einfachen Beweis.

Durch fortgesetzte Polynomdivision erhält man nun:

Satz

Jedes nicht konstante komplexe Polynom läßt sich in Linearfaktoren zerlegen.

Es gibt eine interessante Anwendung auf reelle Polynome:

Satz

Es sei $p(z) = \sum_{i=0}^n a_i z^i$ ein Polynom mit **reellen** Koeffizienten. Dann gilt:

1. Ist $\alpha \in \mathbb{C}$ und $p(\alpha) = 0$, so ist auch $p(\bar{\alpha}) = 0$.
2. Die Anzahl der nicht-reellen Nullstellen von p ist gerade.
3. Ist n ungerade, so besitzt p mindestens eine reelle Nullstelle.
4. Besitzt p keine reelle Nullstelle, so ist $p(z)$ Produkt von Potenzen von quadratischen reellen Polynomen.

BEWEIS: Ist $\alpha \in \mathbb{C}$, so ist $\alpha = \bar{\alpha} \iff \alpha \in \mathbb{R}$.

1) Da alle a_i reell sind, ist

$$p(\bar{\alpha}) = \overline{p(\alpha)} = 0.$$

2) Ist $\alpha \notin \mathbb{R}$, so ist $\bar{\alpha} \neq \alpha$. Also kann man die nicht-reellen Nullstellen in Paaren zusammenfassen.

3) Es gibt $n = 2m + 1$ komplexe Nullstellen. Wegen (2) muß mindestens eine davon reell sein.

4) Sei $\alpha = a + bj$ eine nicht-reelle Nullstelle. Dann enthält $p(z)$ den Faktor

$$(z - \alpha)(z - \bar{\alpha}) = (z - a - bj)(z - a + bj) = (z - a)^2 - (bj)^2 = z^2 - 2az + (a^2 + b^2),$$

und der ist ein quadratisches Polynom mit reellen Koeffizienten. Fortgesetzte Polynomdivision ergibt die Behauptung. ■