

**Einführung in die
Analytische Geometrie und Lineare Algebra**

für Studierende des Köln-Kollegs
im Leistungskurs des 5. Semesters

Unterrichtsbegleitende Skripten sowie
Übungen mit ausführlichen Lösungen

Norbert Klingen

Köln 2001/2002

Inhalt

§1 Von der Anschauung zum Vektorraumbegriff	1
a. Pfeile und Vektoren.....	1
b. Koordinaten von Vektoren.....	1
c. Vektorsumme und skalare Multiplikation.....	2
d. Der Vektorraumbegriff.....	3
e. Unterräume und Linearkombinationen.....	5
f. Lineare Unabhängigkeit und Koeffizientenvergleich.....	7
g. Basen und Dimension.....	8
§2 Punkte, Geraden, Ebenen, affine Räume	10
a. Punkte und Ortsvektoren.....	10
b. Geraden.....	10
c. Ebenen.....	11
d. Affine Koordinatensysteme.....	13
§3 Lineare Gleichungssysteme und das Gauß'sche Eliminationsverfahren	13
a. Standardform.....	13
b. Matrixschreibweise.....	14
c. Gauß'sches Eliminationsverfahren.....	14
d. Der Rang.....	14
e. Lösbarkeit.....	14
f. Lösungsmenge.....	15
g. Homogene Gleichungssysteme.....	15
h. Rang und lineare Unabhängigkeit.....	15
i. Rang und Erzeugendensysteme.....	16
j. Koordinatengleichungen affiner Mengen.....	17
§4 Länge, Orthogonalität, Winkel – Skalarprodukt	19
a. Längen.....	19
b. Skalarprodukt.....	20
c. Orthogonalität.....	21
d. Winkel.....	22
e. Normalenvektoren und Koordinatengleichungen.....	23
f. Abstände.....	24

§1 Von der Anschauung zum Vektorraumbegriff

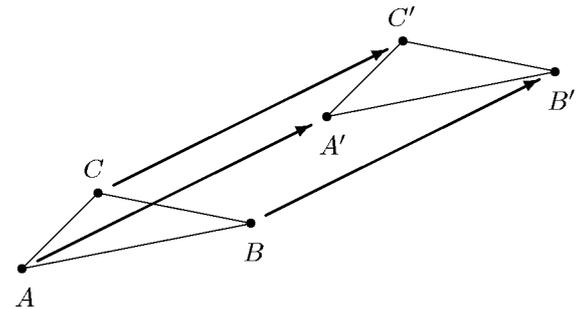
a. Pfeile und Vektoren. Die in der Physik in den verschiedensten Formen auftretenden *Vektoren* (Kräfte, Geschwindigkeiten, Feldstärken, etc.) erscheinen dort in der Regel als sogenannte *Pfeile*.

Ein Pfeil wird durch zwei Punkte bestimmt, den Anfangs- und den Endpunkt.

Innerhalb der Mathematik kommen solche Pfeile als *Verschiebungspfeile* vor: Dabei versteht man unter einer Verschiebung eine Abbildung der gesamten Anschauungsebene oder des Anschauungsraumes, bei der *jeder Punkt* in einer festen Richtung und um einen festen Betrag verschoben wird.

Nun ist aber wiederum aus der Physik bekannt, dass es bei den dort auftretenden ‘vektoriellen’ Größen nicht auf den Anfangspunkt des Pfeils ankommt. Vektorielle Größen können durch verschiedene Pfeile dargestellt werden, wenn diese nur dieselbe Länge, Richtung und Orientierung haben.

Dies kann man auch innerhalb der Mathematik an den Verschiebungspfeilen verdeutlichen. Wir sehen an der nebenstehenden Skizze: Dieselbe Verschiebung kann durch viele verschiedene Pfeile beschrieben werden. Alle diese Pfeile stimmen aber in *Länge, Richtung und Orientierung* überein. Man nennt solche Pfeile *vektorgleich* und sagt, sie bestimmen denselben Vektor:



Zwei Pfeile heißen *vektorgleich*, wenn sie dieselbe Verschiebung der Ebene bzw. des Raumes beschreiben; dies ist genau dann der Fall, wenn sie in Länge, Richtung und Orientierung übereinstimmen.

In der obigen Skizze sind die eingezeichneten Pfeile vektorgleich, wir sagen auch: sie bestimmen denselben Vektor. Wir schreiben dafür:

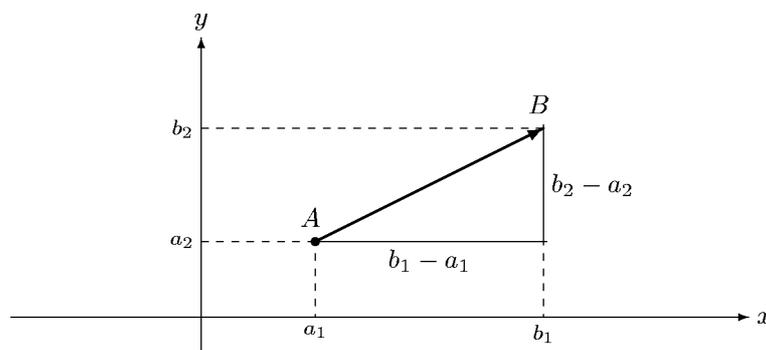
$$\overrightarrow{AA'} = \overrightarrow{BB'} = \overrightarrow{CC'}.$$

Das Symbol $\overrightarrow{AA'}$ bezeichnet also den *Vektor*, der durch den Pfeil von A nach A' bestimmt wird.

Zur Übung: Machen Sie sich an obiger Skizze klar, dass auch die folgenden Vektoren übereinstimmen:

$$\overrightarrow{AB} = \overrightarrow{A'B'}, \quad \overrightarrow{AC} = \overrightarrow{A'C'}, \quad \overrightarrow{BC} = \overrightarrow{B'C'}.$$

b. Koordinaten von Vektoren. Wir wollen nun Vektoren nicht geometrisch, sondern rechnerisch behandeln. Punkte in der Ebene oder im Raum werden durch ihre Koordinaten beschrieben, etwa $A = (a_1, a_2)$, $B = (b_1, b_2)$. Wie kann man nun den Vektor \overrightarrow{AB} beschreiben? Nach Definition ist der Vektor durch Länge, Richtung und Orientierung festgelegt. Hierbei ist die Angabe der Richtung rechnerisch nicht so leicht zu behandeln wie die folgende Überlegung: Man ergänzt einen Pfeil, der in einem Koordinatensystem gegeben ist, durch achsenparallele



Katheten zu einem rechtwinkligen Dreieck (siehe Skizze). Richtung und Länge des Vektors, der die Hypotenuse des rechtwinkligen Dreiecks bildet, sind dann durch die Katheten festgelegt. Dies ist ähnlich wie bei der Definition des Anstiegs einer Geraden. Aber anders als dort kommt es auf *beide* Katheten und nicht nur ihr Verhältnis an, und überdies ist wegen der Orientierung auch die *Reihenfolge* der Punkte und damit das *Vorzeichen* der Werte $b_2 - a_2$ bzw. $b_1 - a_1$ wichtig.

Sind zwei Punkte $A = (a_1, a_2)$ und $B = (b_1, b_2)$ in einem Koordinatenkreuz gegeben, so ist der Vektor \overrightarrow{AB} eindeutig festgelegt durch das Zahlenpaar $\begin{pmatrix} b_1 - a_1 \\ b_2 - a_2 \end{pmatrix}$. Man nennt dies die *Koordinaten des Vektors*. Sie werden spaltenweise geschrieben, um sie von Punktkoordinaten zu unterscheiden.

Die gleichen Überlegungen gelten auch für Punkte im Raum mit 3 Koordinaten:

$$A = (a_1, a_2, a_3), B = (b_1, b_2, b_3) \implies \overrightarrow{AB} = \begin{pmatrix} b_1 - a_1 \\ b_2 - a_2 \\ b_3 - a_3 \end{pmatrix}.$$

Für einen beliebigen Punkt A definiert man den sog. *Ortsvektor* als den Verbindungsvektor \overrightarrow{OA} vom Koordinatenursprung $O = (0, 0, 0)$ zum Punkt A . Wir erhalten:

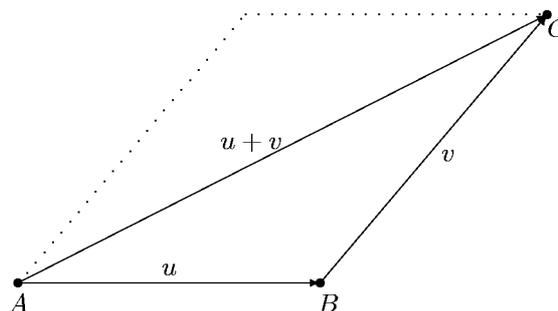
$$A = (a_1, a_2, a_3) \implies \overrightarrow{OA} = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix}.$$

Damit haben Punkt und Ortsvektor dieselben Koordinaten.

c. Vektorsumme und skalare Multiplikation. Ein in der Physik wichtiges Konzept ist die Addition von Vektoren. Anschaulich lässt sie sich folgendermaßen beschreiben:

Zwei Vektoren u und v werden addiert, indem man beide als Pfeile darstellt, und zwar so, dass der Endpunkt des ersten gerade der Anfangspunkt des zweiten ist: $u = \overrightarrow{AB}$, $v = \overrightarrow{BC}$ (man 'setzt v am Ende von u an'). Dann ist der Summenvektor gegeben als Verbindungsvektor vom Anfangspunkt des ersten zum Endpunkt des zweiten Pfeils:

$$\overrightarrow{AB} + \overrightarrow{BC} = \overrightarrow{AC}$$



Deutet man die Vektoren als Verschiebungen (der Ebene oder des Raumes), so kann man die Vektorsumme besonders einfach beschreiben. Verschiebt man A zuerst in Richtung u , so kommt man zum Punkt B ; verschiebt danach in Richtung v , so kommt man zum Punkt C . Insgesamt ergibt sich so die Verschiebung (von A nach C) um den Vektor $u + v$:

Die Vektorsumme $u + v$ ist der Verschiebungsvektor zur 'Hintereinanderausführung' der beiden Verschiebungen zu u und v .

Schließlich kann man die Vektorsumme mit Hilfe der Koordinaten besonders einfach berechnen:

$$u = \begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix}, v = \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} \implies u + v = \begin{pmatrix} u_1 + v_1 \\ u_2 + v_2 \\ u_3 + v_3 \end{pmatrix}$$

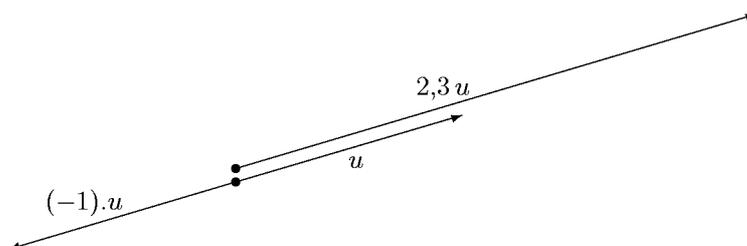
Man erhält die einzelnen Koordinate des Summenvektors $u + v$ als Summe der entsprechenden Koordinaten von u und v .

Neben der Summe gibt es auch eine Multiplikation, jedoch werden nicht zwei Vektoren multipliziert, sondern ein Vektor mit einer reellen Zahl (*skalare Multiplikation*): Ist v ein Vektor und r eine reelle Zahl, so ist $r \cdot v$ wieder ein Vektor.

Dessen Koordinaten erhält man, indem man alle Koordinaten von v mit derselben Zahl r multipliziert:

$$v = \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} \implies r \cdot v = \begin{pmatrix} r v_1 \\ r v_2 \\ r v_3 \end{pmatrix}.$$

Anschaulich bedeutet dies für $r > 0$ eine Änderung der Länge des Vektors um den Faktor r ohne Richtungs- und Orientierungsänderung. Ist r negativ, so wird zusätzlich die Orientierung umgekehrt.



d. Der Vektorraumbegriff. Für diese Vektorsumme und skalare Multiplikation gelten nun eine Reihe von mehr oder minder einsichtigen Eigenschaften, die wir zunächst einmal auflisten. Um diese einfacher formulieren zu können, bezeichnen wir mit V die Menge aller Vektoren (also je nach betrachteter Situation die Menge aller Vektoren der Ebene oder des Raumes, oder alle Verschiebungen bzw. deren Koordinaten). Es gelten die folgenden Eigenschaften:

Zunächst für die Addition:

(S) (Vektorsumme)

$$u, v \in V \implies u + v \in V.$$

(K) (Kommutativgesetz)

$$u + v = v + u \text{ für alle } u, v \in V.$$

(A) (Assoziativgesetz)

$$(u + v) + w = u + (v + w) \text{ für alle } u, v, w \in V.$$

(N) (Existenz eines Nullvektors)

Es gibt ein $o \in V$ (genannt 'Nullvektor') mit der Eigenschaft:

$$u + o = u \text{ für alle } u \in V.$$

(G) (Existenz eines Gegenvektors)

Zu jedem Vektor $v \in V$ gibt es ein $-v \in V$ (genannt 'Gegenvektor' von v) mit der Eigenschaft:

$$v + (-v) = o \text{ für alle } v \in V.$$

Und nun für die skalare Multiplikation:

(SM) (Skalare Multiplikation)

$$u \in V, r \in \mathbb{R} \implies r \cdot u \in V.$$

(A) (Assoziativgesetz)

$$r \cdot (s \cdot u) = (rs) \cdot u \text{ für alle } u \in V, r, s \in \mathbb{R}.$$

(E) (Gesetz der Eins)

$$1 \cdot u = u \text{ für alle } u \in V.$$

Schließlich als Koppelung für Addition und skalare Multiplikation:

(D1) ('rechtes' Distributivgesetz)

$$r \cdot (u + v) = r \cdot u + r \cdot v \text{ für alle } u, v \in V, r \in \mathbb{R}.$$

(D2) ('linkes' Distributivgesetz)

$$(r + s) \cdot u = r \cdot u + s \cdot u \text{ für alle } u \in V, r, s \in \mathbb{R}.$$

Viele Überlegungen und Untersuchungen über Vektoren sind nun nicht von den konkreten Gegebenheiten (Anschauungsraum, Verschiebungsvektoren, Koordinatenbeschreibung) abhängig, vielmehr beruhen sie nur auf den obigen Gesetzmäßigkeiten. Man abstrahiert nun von den konkreten Gegebenheiten und nennt *jede* Menge V mit den obigen Eigenschaften einen *Vektorraum*. So kommt man zu der folgenden allgemeinen Definition:

Definition: Ein *Vektorraum* ist eine Menge V , auf der eine Addition ' $u + v$ ' (für je zwei Vektoren $u, v \in V$) und eine skalare Multiplikation ' $r \cdot u$ ' (für beliebige Vektoren $u \in V$ und reelle Zahlen $r \in \mathbb{R}$) definiert ist, so dass alle oben genannten Eigenschaften erfüllt sind.

Beispiele: a) Der Raum der Verschiebungen des 2- und 3-dimensionalen Punktraumes.
b) Die Menge der Verbindungsvektoren des 2- und des 3-dimensionalen Punktraumes.
c) Die Menge der Tripel reeller Zahlen

$$\mathbb{R}^3 = \left\{ \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} \mid a_1, a_2, a_3 \in \mathbb{R} \right\}$$

wie sie als Koordinaten auftraten.

d) Dieses letzte Beispiel kann man nun unmittelbar verallgemeinern zu den Vektorräumen

$$\mathbb{R}^n = \left\{ \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} \mid a_1, \dots, a_n \in \mathbb{R} \right\}$$

der ' n -Tupel' reeller Zahlen (n eine beliebige natürliche Zahl). Dabei sind Addition und skalare Multiplikation wie bei den Koordinaten 2- bzw. 3-dimensionaler Vektoren *komponentenweise* erklärt:

$$\begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} + \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix} = \begin{pmatrix} a_1 + b_1 \\ \vdots \\ a_n + b_n \end{pmatrix}, \quad r \cdot \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} = \begin{pmatrix} ra_1 \\ \vdots \\ ra_n \end{pmatrix}.$$

e) Eng damit zusammenhängend sind die sog. Matrizenvektorräume. *Matrizen* sind rechteckige Zahlenschemata, z. B.

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

Dieses M ist eine Matrix mit 4 Zeilen und 3 Spalten. Allgemein ist eine beliebige Kombination von Zeilen- und Spaltenzahl möglich.

Übung: Überlegen Sie sich, wie man auf der Menge aller Matrizen mit einer festen Zeilenzahl m und Spaltenzahl n eine Addition und skalare Multiplikation definieren sollte, so dass ein Vektorraum entsteht. Dieser wird mit dem Symbol $\mathbb{R}^{m \times n}$ bezeichnet.

f) Eine weitere Serie von Beispielen, die nicht von der Anschauung und Geometrie herkommen, sind die *Funktionsräume*:

Satz: Ist D eine Menge und V die Menge aller Funktionen $f : D \rightarrow \mathbb{R}$, die jedem $x \in D$ ($D = \text{Definitionsbereich}$) eine reelle Zahl $f(x)$ zuordnen, so bildet V einen Vektorraum, wenn man Addition und skalare Multiplikation wertweise definiert:

A) $f, g \in V \implies f + g \in V$, definiert durch

$$(f + g)(x) = f(x) + g(x) \quad \text{für alle } x \in D.$$

B) $f \in V, r \in \mathbb{R} \implies r \cdot f \in V$, definiert durch

$$(r \cdot f)(x) = r \cdot f(x) \quad \text{für alle } x \in D.$$

Zur *Begründung* muss man die Axiome eines Vektorraums überprüfen. Zunächst einmal sind die so definierte Summe $f + g$ und Vielfachen $r \cdot f$ wieder Funktionen von D in \mathbb{R} , gehören also zu V . Damit sind die Forderungen (S) der Vektorsumme und (SM) der skalaren Multiplikation (siehe oben) erfüllt.

Da sich die eben definierten Verknüpfungen $f + g$ bzw. $r \cdot f$ unmittelbar auf das Rechnen mit reellen Zahlen zurückführen: $f(x) + g(x)$ bzw. $r \cdot f(x)$, übertragen sich die bekannten Rechengesetze von \mathbb{R} unmittelbar auf V :

die beiden Assoziativgesetze (A), das Kommutativgesetz (K), die beiden Distributivgesetze (D1), (D2), die Existenz von Null (N) und Gegenvektor (G) sowie das Gesetz der Eins (E). Die Nullfunktion ist dabei die Funktion, die an jeder Stelle $x \in D$ den Wert 0 hat: $f(x) = 0$ für alle $x \in D$, und die Gegenfunktion $-f$ ist gegeben durch $(-f)(x) = -f(x)$, d. h. durch wertweise Vorzeichenumkehr.

Die Menge \mathbb{R}^D aller reellwertigen Funktionen $f : D \rightarrow \mathbb{R}$ mit beliebigem Definitionsbereich D bildet bei wertweiser Addition und skalarer Multiplikation einen Vektorraum.

e. Unterräume und Linearkombinationen. Viele weitere Beispiele von Vektorräumen entstehen, wenn man in einem bereits gegebenen Vektorraum V kleinere sog. *Unterräume* U konstruiert. Für diese ist der Nachweis der Vektorraum-Eigenschaften vereinfacht, da die in V gültigen Axiome (K), (A), (D1), (D2) natürlich erst recht im kleineren Bereich U gelten. Für die Existenz (N) der Null muss man lediglich überprüfen, dass der in V ja vorhandene Nullvektor o zu U gehört! Ebenso sind Addition und skalare Multiplikation bereits von V her definiert, so dass man lediglich (S) und (SM) überprüfen muss, d. h. dass Summe $u + u'$ und skalares Vielfaches $r \cdot u$ zu U gehören, wenn u, u' zu U gehören.

Dies hat dann wegen der Eigenschaft $-u = (-1) \cdot u$ automatisch zur Folge, dass für jedes $u \in U$ der Gegenvektor $-u$ zu U gehört, so dass (G) dann auch für U gilt.

Um also eine Teilmenge $U \subset V$ als Unterraum des Vektorraums V nachzuweisen, muss man nur die folgenden 3 Eigenschaften überprüfen:

(N')	$o \in U,$
(S)	$u, u' \in U \implies u + u' \in U,$
(SM)	$u \in U \implies r \cdot u \in U$ für alle $r \in \mathbb{R}.$

Übung: Zeigen Sie:

a) $U = \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \mid 3x - 2y + z = 0 \right\}$ ist ein Unterraum des \mathbb{R}^3 .

b) $U = \left\{ \begin{pmatrix} r_1 \\ r_2 \\ 0 \end{pmatrix} \mid r_1, r_2 \in \mathbb{R} \right\}$ ist ein Unterraum des \mathbb{R}^3 . Bilden Sie ähnliche Beispiele.

c) $C = \left\{ \begin{pmatrix} a & b \\ -b & a \end{pmatrix} \mid a, b \in \mathbb{R} \right\}$ ist ein Unterraum des Raumes $\mathbb{R}^{2 \times 2}$ aller reellen 2×2 -Matrizen.

d) Für Teilmengen $I \subset \mathbb{R}$ ist die Menge aller stetigen Funktionen auf I

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

ein Unterraum im Vektorraum $V = \mathbb{R}^I$ aller Funktionen $I \rightarrow \mathbb{R}$.

e) Gleiches gilt, wenn man *stetig* durch *differenzierbar* oder *integrierbar* ersetzt.

f) Die Menge aller ganzrationalen Funktionen ist ein Unterraum des Vektorraums $\mathbb{R}^{\mathbb{R}}$ aller Funktionen $\mathbb{R} \rightarrow \mathbb{R}$.

Bei a) ist wichtig, dass auf der rechten Seite der Wert 0 steht. Ebenso bei b), dass dort die dritte Koordinate 0 (und kein anderer Wert) ist. Wo liegen die Probleme bei a) und b), wenn man statt der 0 eine beliebige andere Zahl (etwa 1) an der entsprechenden Stelle vorschreibt? Die Gültigkeit von d) beruht im wesentlichen auf der Summen- bzw. Faktorregel: Summen stetiger Funktionen sind wieder stetig bzw. konstante Vielfache stetiger Funktionen sind wieder stetig. Gleiches gilt für *differenzierbar* und *integrierbar*. Gilt dies auch für *monoton steigend* oder für einfach nur *monoton*?

Abschließend sei noch das folgende Verfahren erwähnt, mit dem man von festen Vektoren ausgehend systematisch Unterräume erzeugen kann. Ist $v \in V$ fest vorgegeben, so bildet die Menge aller Vielfachen von v einen Unterraum von V :

$$\mathbb{R}v = \{rv \mid r \in \mathbb{R}\} \text{ ist ein Unterraum von } V.$$

Zum Nachweis beachte man $0v = o \in \mathbb{R}v$, $rv + sv = (r + s)v \in \mathbb{R}v$ und schließlich $s(rv) = (sr)v \in \mathbb{R}v$. Damit sind die drei Unterraumkriterien erfüllt.

Statt der Vielfachen *eines* Vektors kann man auch von *mehreren* Vektoren v_1, \dots, v_k ausgehend deren *Linearkombinationen* betrachten. Darunter versteht man Summen von Vielfachen dieser Vektoren:

Linearkombinationen von Vektoren v_1, \dots, v_k sind Terme der Form

$$r_1v_1 + \dots + r_kv_k = \sum_{i=1}^k r_iv_i \quad (\text{mit Koeffizienten } r_1, \dots, r_k \in \mathbb{R})$$

Die Gesamtheit aller Linearkombination von fest vorgegebenen Vektoren v_1, \dots, v_k ($k \in \mathbb{N}$ beliebig) bildet ebenfalls einen Unterraum von V :

$$\mathbb{R}v_1 + \mathbb{R}v_2 + \dots + \mathbb{R}v_k = \left\{ \sum_{i=1}^k r_iv_i \mid r_i \in \mathbb{R} \right\} \text{ ist ein Unterraum von } V.$$

Es ist dies zugleich auch der kleinste Unterraum von V , in dem die vorgegebenen Vektoren v_1, \dots, v_k enthalten sind, denn wenn erst v_1, \dots, v_k dazugehören, müssen auch deren Vielfache und von diesen wiederum sämtliche Summen dazugehören. Man sagt, der Unterraum $\mathbb{R}v_1 + \dots + \mathbb{R}v_k$ wird von v_1, \dots, v_k erzeugt.

Beispiele: Wir betrachten die Vektoren

$$u = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}, \quad v = \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}, \quad e_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad e_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \quad e_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}.$$

Dann gilt:

$$\mathbb{R}u = \left\{ \begin{pmatrix} r \\ 2r \\ 3r \end{pmatrix} \mid r \in \mathbb{R} \right\}, \quad \mathbb{R}v = \left\{ \begin{pmatrix} r \\ -r \\ 0 \end{pmatrix} \mid r \in \mathbb{R} \right\} = \left\{ \begin{pmatrix} s \\ -s \\ 0 \end{pmatrix} \mid s \in \mathbb{R} \right\},$$

$$\mathbb{R}u + \mathbb{R}v = \left\{ ru + sv = \begin{pmatrix} r + s \\ 2r - s \\ 3r \end{pmatrix} \mid r, s \in \mathbb{R} \right\},$$

$$\mathbb{R}e_1 + \mathbb{R}e_3 = \left\{ re_1 + se_3 = \begin{pmatrix} r \\ 0 \\ s \end{pmatrix} \mid r, s \in \mathbb{R} \right\}.$$

f. Lineare Unabhängigkeit und Koeffizientenvergleich. Die geometrische Bedeutung von Linearkombinationen haben wir an diversen Beispielen aus dem Lehrbuch veranschaulicht. Etwa Lehrbuch, S. 40, Aufgabe 4: Die diversen Vektoren (Seitenhalbierende, Mittellinien) lassen sich alle als Linearkombinationen der drei Seitenvektoren $\vec{a}, \vec{b}, \vec{c}$ darstellen. Wir haben in diesem Beispiel aber auch gesehen, dass es durchaus *mehrere* Linearkombinationen geben kann, die *denselben* Vektor darstellen. Als Ursache dieser Tatsache haben wir die *lineare Relation* $\vec{a} + \vec{b} + \vec{c} = \vec{o}$ zwischen den drei Vektoren festgestellt.

Unter Verwendung dieser Relation haben wir dann überall \vec{c} durch $-\vec{a} - \vec{b}$ ersetzt und so alle Vektoren als Linearkombination nur von \vec{a} und \vec{b} dargestellt. Anders als bei der Darstellung durch die drei Vektoren $\vec{a}, \vec{b}, \vec{c}$ ergab sich hier, dass jeder Vektor nur *eine* Darstellung als Linearkombination von \vec{a}, \vec{b} besaß. Ursache dafür war, dass zwischen \vec{a} und \vec{b} eben *keine* lineare Relation mehr bestand, dass sie *linear unabhängig* waren.

Ein System von Vektoren v_1, \dots, v_k nennt man *linear abhängig*, wenn sich (wenigstens) einer der Vektoren als Linearkombination der *übrigen* darstellen lässt. Andernfalls nennt man die Vektoren *linear unabhängig*.

Bei 2 Vektoren v_1, v_2 bedeutet lineare Abhängigkeit, dass einer der Vektoren ein Vielfaches des anderen sein muss. Dies ist bei konkret gegebenen Vektoren einfach zu überprüfen (siehe Übung (2), Aufgabe 5).

Bei 3 und mehr Vektoren ist die Überprüfung schon mühseliger. Man benutzt dann in der Regel den nachfolgenden Satz, der die entscheidenden Charakterisierungen der linearen Unabhängigkeit enthält:

Satz. Für k Vektoren v_1, \dots, v_k in einem Vektorraum V sind die folgenden Aussagen äquivalent:

i) (Lineare Unabhängigkeit)

Keiner der Vektoren v_i ist eine Linearkombination der *übrigen*.

ii) (Nur die triviale Nulldarstellung bzw. Keine echte lineare Relation)

Stellt eine Linearkombination den Nullvektor o dar, so müssen *alle* Koeffizienten 0 sein:

$$r_1v_1 + r_2v_2 + \dots + r_kv_k = o \implies r_1 = r_2 = \dots = r_k = 0.$$

iii) (Koeffizientenvergleich) Stellen zwei Linearkombinationen denselben Vektor dar, so müssen die einander entsprechenden Koeffizienten übereinstimmen:

$$r_1v_1 + \dots + r_kv_k = s_1v_1 + \dots + s_kv_k \iff r_1 = s_1, \dots, r_k = s_k.$$

Die Eigenschaft ii) benutzt man meist zum Nachweis der linearen Unabhängigkeit, während iii) die wichtigste Anwendung beinhaltet. Die Namensgebung in iii) spricht hoffentlich für sich: Stellen zwei (formal verschiedene) Linearkombinationen von k Vektoren v_1, \dots, v_k denselben Vektor dar, so ist dies bei *linear unabhängigen* Vektoren v_1, \dots, v_k nur möglich, wenn jeweils die einander entsprechenden Koeffizienten übereinstimmen. Auf diese Weise ist eine Vektorgleichung äquivalent zu k Gleichungen für Zahlen.

Die Eigenschaft ii) ist folgendermaßen zu verstehen: Aus k beliebigen Vektoren v_1, \dots, v_k kann man immer eine Linearkombination bilden, die den Nullvektor o darstellt; man wählt alle Koeffizienten $r_i = 0$:

$$0 \cdot v_1 + \dots + 0 \cdot v_k = o.$$

Man nennt dies die ‘triviale’ (offensichtliche) Darstellung des Nullvektors. ii) besagt nun, dass bei linear unabhängigen Vektoren *nur die triviale Nulldarstellung* möglich ist.

Eine andere Deutung von Bedingung ii) ist die folgende: Eine Gleichung der Form $r_1 v_1 + \dots + r_k v_k = o$ stellt eine echte *Relation* (Beziehung) zwischen den Vektoren dar, wenn nicht sämtliche $r_i = 0$ sind. Man kann also sagen, dass k Vektoren linear unabhängig sind, wenn zwischen ihnen *keine echte Relation* besteht.

Beweis des Satzes: i) \implies ii): Wir setzen i) voraus und nehmen an, ii) wäre falsch. Das bedeutet, dass doch eine *nicht-triviale Nulldarstellung existiert*: $o = r_1 v_1 + \dots + r_k v_k$. Wenn dies nicht die triviale Nulldarstellung ist, muss wenigstens einer der Koeffizienten $\neq 0$ sein, etwa $r_1 \neq 0$. Dann kann man diese Gleichung aber nach v_1 auflösen:

$$o = r_1 v_1 + \dots + r_k v_k \iff r_1 v_1 = -r_2 v_2 - \dots - r_k v_k \iff v_1 = -\frac{r_2}{r_1} v_2 - \dots - \frac{r_k}{r_1} v_k.$$

(Dabei wird wesentlich benutzt, dass $r_1 \neq 0$ ist.) Damit wäre aber v_1 eine Linearkombination der übrigen Vektoren v_2, \dots, v_k , was nach i) nicht sein kann.

Genauso folgert man einen Widerspruch, wenn ein anderer Koeffizient $r_i \neq 0$ ist.

ii) \implies iii): Wir brauchen nur die Richtung \implies nachzuweisen, da die Umkehrung selbstverständlich ist. Dann erhalten wir unter Verwendung von ii):

$$\begin{aligned} r_1 v_1 + \dots + r_k v_k = s_1 v_1 + \dots + s_k v_k &\implies (r_1 - s_1) v_1 + \dots + (r_k - s_k) v_k = o \\ &\stackrel{ii)}{\implies} r_1 - s_1 = \dots = r_k - s_k = 0 \\ &\implies r_1 = s_1, \dots, r_k = s_k \end{aligned}$$

Damit ist iii) bewiesen.

iii) \implies i): Wäre i) falsch, etwa v_1 eine Linearkombination der übrigen:

$$v_1 = r_2 v_2 + \dots + r_k v_k,$$

so erhielte man zwei verschiedene Linearkombinationen für denselben Vektor:

$$1 \cdot v_1 + 0 \cdot v_2 + \dots + 0 \cdot v_k = 0 \cdot v_1 + r_2 v_2 + \dots + r_k v_k.$$

Dies ist aber ein Widerspruch zu iii). Also kann v_1 keine Linearkombination der anderen Vektoren sein. Genauso schließt man, wenn ein anderes v_i eine Linearkombination der übrigen Vektoren wäre.

Beispiele für die Anwendung des Koeffizientenvergleichs zum Beweis allgemeiner geometrischer Sachverhalte siehe Übung (3).

g. Basen und Dimension. Im Zusammenhang mit Unterräumen haben wir bereits von erzeugenden Vektoren gesprochen. Wir definieren:

Unter einem *Erzeugendensystem* eines Vektorraumes V versteht man ein System von Vektoren $v_1, \dots, v_n \in V$ mit der Eigenschaft:
Jeder Vektor $v \in V$ ist eine Linearkombination von v_1, \dots, v_n :

$$v = \sum_{i=1}^n r_i v_i \quad \text{mit geeigneten Koeffizienten } r_1, \dots, r_n \in \mathbb{R}.$$

Ist nun das Erzeugendensystem v_1, \dots, v_n zusätzlich auch linear unabhängig, so sind die Koeffizienten r_1, \dots, r_n durch v *eindeutig* bestimmt. Wir definieren:

Unter einer *Basis* eines Vektorraumes V versteht man ein linear unabhängiges Erzeugendensystem $v_1, \dots, v_n \in V$. Dies bedeutet:
Jeder Vektor $v \in V$ besitzt eine eindeutige Darstellung als Linearkombination der Basisvektoren v_1, \dots, v_n :

$$v = \sum_{i=1}^n r_i v_i \quad \text{mit eindeutig bestimmten } r_1, \dots, r_n \in \mathbb{R}.$$

Fixiert man also eine Basis v_1, \dots, v_n eines Vektorraumes, so wird jeder Vektor v des abstrakten Vektorraumes in eindeutiger Weise durch n reelle Zahlen r_1, \dots, r_n beschrieben. Man fasst diese

reellen Zahlen zu einem Vektor $\begin{pmatrix} r_1 \\ \vdots \\ r_n \end{pmatrix} \in \mathbb{R}^n$ zusammen, dem sog. *Koordinatenvektor von v*

bzgl. der Basis v_1, \dots, v_n . Dieser Name kommt von folgendem fundamentalen Beispiel her:

Die Vektoren

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

bilden eine Basis, die sog. kanonische Basis des \mathbb{R}^n , und es gilt:

$$x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = x_1 e_1 + \dots + x_n e_n = \sum_{i=1}^n x_i e_i.$$

Die Koordinaten des Vektors $x \in \mathbb{R}^n$ sind also gerade die Koeffizienten in der Linearkombinationsdarstellung von x durch die Basis e_1, \dots, e_n . Dieses Beispiel motiviert auch die nachfolgende Definition:

Besitzt ein Vektorraum V eine Basis v_1, \dots, v_n aus n Vektoren, so sagt man: V hat die Dimension n .

Fixiert man eine Basis v_1, \dots, v_n eines Vektorraumes V , so erhält man eine umkehrbar eindeutige Zuordnung zwischen den Vektoren v des abstrakten Vektorraumes V und deren Koordinatenvektoren im \mathbb{R}^n :

$$V \ni v = \sum_{i=1}^n r_i v_i \longleftrightarrow \begin{pmatrix} r_1 \\ \vdots \\ r_n \end{pmatrix} \in \mathbb{R}^n.$$

Bei dieser umkehrbar eindeutigen Zuordnung gehen Vektorsumme und skalare Multiplikation von V in die wohlbekanntere Summe und skalare Multiplikation im \mathbb{R}^n über:

$$v \longleftrightarrow \begin{pmatrix} r_1 \\ \vdots \\ r_n \end{pmatrix}, \quad w \longleftrightarrow \begin{pmatrix} s_1 \\ \vdots \\ s_n \end{pmatrix} \implies v + w \longleftrightarrow \begin{pmatrix} r_1 \\ \vdots \\ r_n \end{pmatrix} + \begin{pmatrix} s_1 \\ \vdots \\ s_n \end{pmatrix}, \quad \lambda v \longleftrightarrow \lambda \begin{pmatrix} r_1 \\ \vdots \\ r_n \end{pmatrix}.$$

Eine solche umkehrbar eindeutige Zuordnung zwischen zwei Vektorräumen nennt man *Isomorphismus*; beide Vektorräume sind hinsichtlich der Vektorraumstruktur völlig gleichwertig, sie sind *isomorph*.

§2 Punkte, Geraden, Ebenen, affine Räume

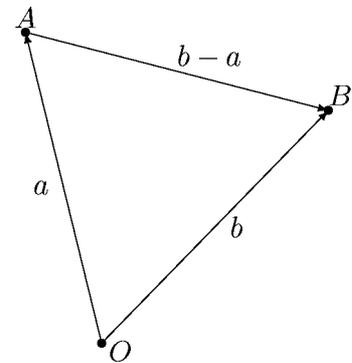
Vektorräume bieten einen guten Rahmen zur Behandlung geometrischer Fragen. Diese werden dabei *analytisch* behandelt, indem man die geometrischen Objekte algebraisch erfasst und dann die Probleme mit den uns vertrauten Mitteln der Algebra löst. Ein wichtiges Mittel zur algebraischen Erfassung geometrischer Sachverhalte ist der Vektorbegriff. Wie wir im ersten Abschnitt gesehen haben, können wir mit Vektoren weitgehend gewohnt ‘rechnen’, genauer, wir können sie addieren und mit Skalaren (Zahlen) multiplizieren und es gelten dafür wohlvertraute Gesetzmäßigkeiten. Insbesondere der Umgang mit Gleichungen und deren äquivalente Umformungen verlaufen nach bekanntem Schema.

Die geometrischen Objekte, die wir im Rahmen der ‘Analytischen Geometrie’ zunächst im Auge haben, sind ‘geradlinig’: Geraden und Ebenen, geradlinig begrenzte ebene und räumliche Körper (Dreiecke, Vierecke, Tetraeder, ...).

a. Punkte und Ortsvektoren. Die genannten geometrischen Objekte sind aus den Grundbausteinen der Geometrie, den ‘Punkten’ aufgebaut. Diese Punkte kann man ebenfalls durch Vektoren erfassen. Dazu fixiert man einen Ausgangs- oder auch Basispunkt O , und erfasst dann alle Punkte X durch den sog. *Ortsvektor* \overrightarrow{OX} . Da Punkte und Ortsvektoren einander eindeutig entsprechen: $X = Y \iff \overrightarrow{OX} = \overrightarrow{OY}$, kann man die geometrischen Fragen vektoriell behandeln. (Man beachte aber, dass der Ortsvektor einen festgewählten Ursprung O erfordert.)

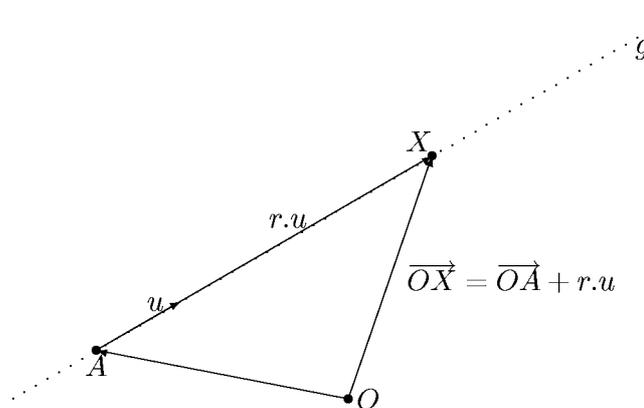
Auch beliebige Verbindungsvektoren zweier Punkte A, B kann man durch die Ortsvektoren $a = \overrightarrow{OA}$ von A und $b = \overrightarrow{OB}$ von B erfassen:

$$\overrightarrow{AB} = \overrightarrow{AO} + \overrightarrow{OB} = \overrightarrow{OB} - \overrightarrow{OA} = b - a.$$



Der Verbindungsvektor zweier Punkte ist also die Differenz aus Ortsvektor des Endpunktes minus Ortsvektor des Anfangspunktes.

b. Geraden. Zur Festlegung einer Geraden braucht man entweder zwei verschiedene Punkte A, B oder einen Punkt A und die Richtung. Die Festlegung der Richtung erfolgt dabei durch einen (Richtungs-) Vektor $u \neq o$. Gehen wir zunächst von dem zweiten Fall aus: Gegeben ein Punkt A und ein Richtungsvektor u ; es bezeichne $g = g(A, u)$ die Gerade durch A mit Richtungsvektor u . Wir wollen sämtliche Punkte X auf der Geraden algebraisch erfassen. Nun liegt ein Punkt X gerade dann auf der Geraden g , wenn der Vektor \overrightarrow{AX} dieselbe Richtung (nicht unbedingt Orientierung) wie der gegebene Richtungsvektor u hat. Das bedeutet, dass der Vektor \overrightarrow{AX} ein Vielfaches von u ist: $\overrightarrow{AX} = r \cdot u$ für ein $r \in \mathbb{R}$. Dabei ist r eine beliebige reelle Zahl; sie



kann auch negativ sein. Wir erhalten so:

$$\boxed{X \in g(A, u) \iff \overrightarrow{AX} = r \cdot u \text{ für ein } r \in \mathbb{R}.} \quad (1)$$

Will man nun direkt den Punkt X erfassen, so löst man die obige Beziehung nach dem Ortsvektor \overrightarrow{OX} auf: $\overrightarrow{AX} = \overrightarrow{OX} - \overrightarrow{OA} = r \cdot u \iff \overrightarrow{OX} = \overrightarrow{OA} + r \cdot u$ und erhält

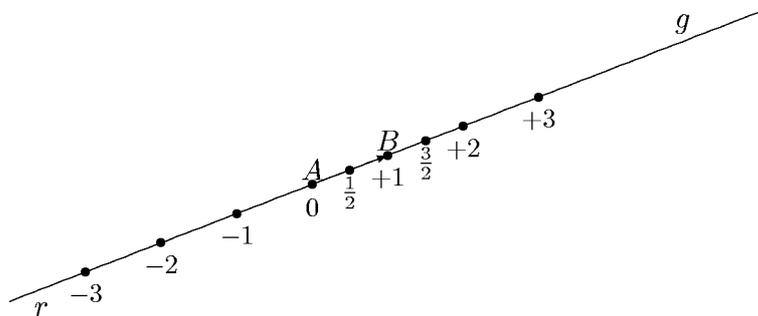
$$\boxed{X \in g(A, u) \iff \overrightarrow{OX} = \overrightarrow{OA} + r \cdot u \text{ für ein } r \in \mathbb{R}.} \quad (1')$$

Man nennt diese beiden Beschreibungen (1) bzw. (1') eine *Parameterdarstellung* der Geraden g durch den Punkt A mit dem Richtungsvektor u . Für jeden Wert des *Parameters* r liefert die Gleichung (1') $\overrightarrow{OX} = \overrightarrow{OA} + r \cdot u$ den Ortsvektor eines Punktes X auf g , und es werden dabei alle Punkte genau einmal erfasst. Es wird also, unabhängig davon, ob die Gerade in einer Ebene oder im Raum verläuft, jeder Punkt X der Geraden g durch *einen* Parameter erfasst.

Ist nun die Gerade durch zwei verschiedene Punkte A, B beschrieben, so wählt man etwa A als den einen Punkt auf g und $u = \overrightarrow{AB}$ als Richtungsvektor ($u \neq o$ da $A \neq B$). Dann erhält man eine Parameterdarstellung für die Gerade $g = g(A, B)$ durch die Punkte $A \neq B$ wie folgt:

$$\boxed{X \in g \iff \overrightarrow{OX} = \overrightarrow{OA} + r \cdot \overrightarrow{AB} \text{ für ein } r \in \mathbb{R}.} \quad (1'')$$

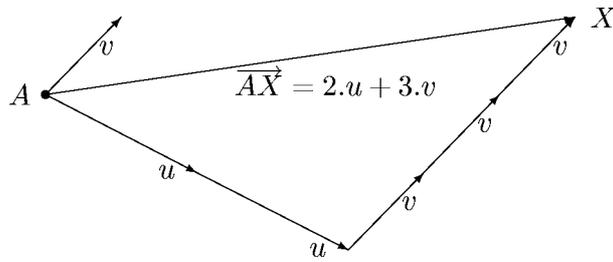
In der nachstehenden Skizze sind bei gegebenen Punkten A, B für einige Geradenpunkte die zugehörigen Parameterwerte angegeben. Man sieht: Der Gerade wird auf diese Weise eine Skala aufgeprägt und jeder Geradenpunkt wird durch einen Zahlwert (Parameter) erfasst. Man benutzt dann zur Rechnung diese Parameterwerte statt der Punkte.



c. Ebenen. Ebenen kann man in ähnlicher Weise wie Geraden durch Parameterdarstellungen beschreiben. Dabei wollen wir unter einer Ebene eine unbegrenzte, nicht gekrümmte Fläche verstehen. Sie setzt sich also aus lauter Geraden zusammen. Zur Festlegung einer Geraden benötigt man zwei verschiedene Punkte A, B . Durch eine Gerade können aber viele Ebenen verlaufen. Um also eine Ebene festzulegen, muss man zu den beiden Punkten A, B noch einen dritten angeben. Dieser darf aber nicht auf der Geraden $g(A, B)$ liegen. Eine Ebene e ist somit festgelegt durch 3 in ihr liegende Punkte A, B, C , die nicht auf einer Geraden liegen, man sagt, die nicht *kollinear* sind.

Wir veranschaulichen diese Situation in der nachfolgenden Skizze, wobei wir zur Vermeidung der Probleme der räumlichen Darstellung die Ebene e in die Zeichenebene gelegt haben. Wir fixieren (wie bei den Geraden) einen Punkt in der Ebene e , etwa A , sowie zwei Richtungsvektoren u und v , die etwa durch die Vektoren $u = \overrightarrow{AB}$ und $v = \overrightarrow{AC}$ gegeben sein können. Die Tatsache, dass die 3 Punkte nicht auf einer Geraden liegen sollen, bedeutet, dass von den beiden Vektoren u, v keiner ein Vielfaches des anderen sein darf, die beiden Vektoren u, v also linear unabhängig sein müssen.

‘Geht’ man nun vom Ausgangspunkt A beliebig weit in Richtung von u und dann in Rich-

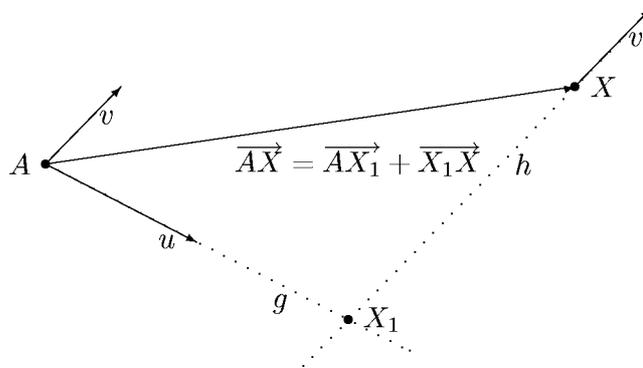


tung von v , so verlässt man die Ebene nicht, man ‘kommt’ immer zu einem Endpunkt X , der in der Ebene liegt. Der dabei ‘zurückgelegte’ Vektor ist von der Form $r.u + s.v$, also eine Linearkombination der Vektoren u, v . Dies bedeutet, dass alle Punkte X mit

$$\overrightarrow{AX} = r.u + s.v \quad (r, s \in \mathbb{R})$$

notwendig zur Ebene gehören.

Aber auch umgekehrt: Liegt X in der Ebene, so lässt sich der Vektor \overrightarrow{AX} stets als Linearkombination von u und v darstellen. In der nachfolgenden Skizze ist verdeutlicht, wie man



die zugehörigen Parameterwerte r, s geometrisch findet. Man betrachte die Gerade g durch A mit Richtung u und die Gerade h durch X mit Richtung v . Diese schneiden sich in einem Punkt X_1 (da sie in einer Ebene liegen, aber nicht parallel sind). Nun ist $\overrightarrow{AX_1}$ ein Vielfaches $r.u$ von u und $\overrightarrow{X_1X}$ ein Vielfaches $s.v$ von v . Also $\overrightarrow{AX} = \overrightarrow{AX_1} + \overrightarrow{X_1X} = r.u + s.v$.

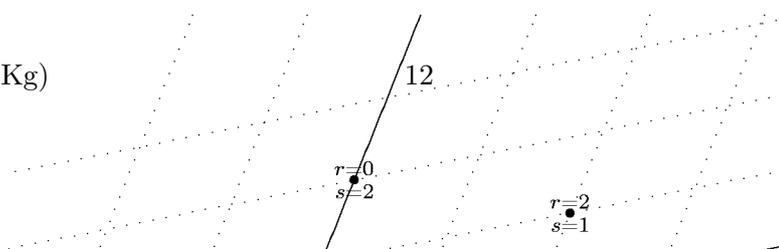
Zusammenfassend erhalten wir so eine Parameterdarstellung für die Ebene $e = e(A, u, v)$, die durch den Punkt A verläuft und die linear unabhängigen Richtungsvektoren u, v hat:

$$\boxed{X \in e(A, u, v) \iff \overrightarrow{AX} = r.u + s.v \text{ für geeignete } r, s \in \mathbb{R}.} \quad (2)$$

Ist e die Ebene durch drei nicht-kollineare Punkte A, B, C : $e = e(A, B, C)$, so nimmt die Parameterdarstellung folgende Gestalt an:

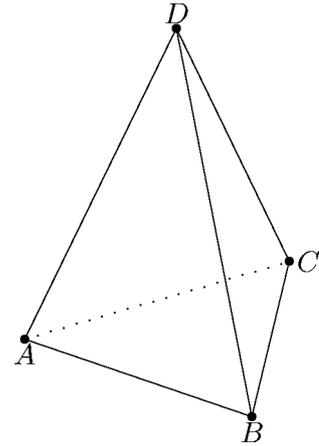
$$\boxed{X \in e(A, B, C) \iff \overrightarrow{OX} = \overrightarrow{OA} + r.\overrightarrow{AB} + s.\overrightarrow{AC} \text{ für geeignete } r, s \in \mathbb{R}.} \quad (2'')$$

In der nachstehenden Skizze sind bei gegebenen Punkten A, B, C die Vektoren $u = \overrightarrow{AB}$, $v = \overrightarrow{AC}$ sowie für einige Punkte der Ebene $e(A, B, C)$ die zugehörigen Parameterwerte r und s angegeben. Man sieht: Jeder Punkt der Ebene wird durch *zwei* Zahlwerte (Parameter r, s) erfasst und man erhält dadurch ein (schiefwinkliges) Koordinatensystem in der Ebene.



d. Affine Koordinatensysteme. Wir wollen nun auch räumliche Gebilde ('Körper') erfassen. Das einfachste darunter ist das Tetraeder. Es hat 4 Eckpunkte $ABCD$; davon bilden zunächst drei Punkte ein Dreieck (etwa ABC) (in nebenstehender Skizze der 'Boden'), während der vierte Punkt (die 'Spitze') *nicht in einer Ebene* mit diesen drei Punkten liegen soll:

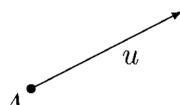
Vier Punkte, die nicht in einer Ebene liegen, bilden ein Tetraeder.



Man erhält so ein pyramidenartiges Gebilde, welches von 4 Dreiecken begrenzt wird. Der Name Tetraeder bedeutet 'Vierflach' oder 'Vierflächner', von vier Flächen begrenzt. Um nun ein solches räumliches Gebilde zu erfassen, braucht man neben einem Ausgangspunkt (etwa A) und den beiden Vektoren $u = \overrightarrow{AB}$, $v = \overrightarrow{AC}$, die die 'Boden'ebene beschreiben, noch einen dritten Vektor: $w = \overrightarrow{AD}$. Da D nicht in einer Ebene mit ABC liegen soll, darf der Vektor $w = \overrightarrow{AD}$ keine Linearkombination von u, v sein, die drei Vektoren u, v, w müssen linear unabhängig sein.

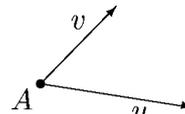
Zur Beschreibung von Geraden und Ebenen haben wir Parameterdarstellungen benutzt. Diese wiederum wurden festgelegt durch Auswahl eines sog. *affinen Koordinatensystems*:

Bei Geraden:



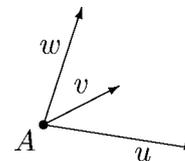
Ein Punkt A und ein Vektor $u \neq 0$.

Bei Ebenen:



Ein Punkt A und zwei Vektoren unterschiedlicher Richtung, d. h. zwei linear unabhängige Vektoren u, v .

Und um den gesamten (dreidimensionalen) *Raum* zu erfassen, benötigt man wiederum einen Punkt A sowie nun *drei* Vektoren u, v, w , die wir wieder mit Anfangspunkt A darstellen. Damit alle Punkte des Raumes erreicht werden, dürfen die Vektoren aber *nicht in einer Ebene liegen*. Dies bedeutet, dass nicht einer der Vektoren als Linearkombination der beiden anderen dargestellt werden kann, d. h. u, v, w müssen linear unabhängig sein.



§3 Lineare Gleichungssysteme und das Gauß'sche Eliminationsverfahren

Alle bisher behandelten geometrischen Fragen führten letztendlich auf *lineare Gleichungssysteme*. Solche Gleichungssysteme treten jedoch nicht nur bei geometrischen Fragen sondern in *allen* Bereichen der Mathematik und in *allen* Anwendungsgebieten auf, so dass die Beherrschung der Lösungsmethoden und der grundlegenden Zusammenhänge der Theorie linearer Gleichungssysteme unabdingbar ist.

a. Standardform. Ein lineares Gleichungssystem kann immer (äquivalent) in die folgende Form gebracht werden:

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

Dabei ist m die Zahl der *Gleichungen* und n die Zahl der *Unbekannten* x_1, \dots, x_n . Die Koeffizienten (Vorfaktoren) a_{ij} der Unbestimmten sind so bezeichnet, dass man an den Indizes i und j erkennen kann, um welche Gleichung es sich handelt (erster Index i) und zu welcher Unbekannten der Koeffizient gehört (Index j).

b. Matrixschreibweise. Man schreibt ein solches Gleichungssystem aus rechentechnischen Gründen in einer verkürzten Matrixschreibweise, indem man die Unbekannten in ihrer Reihenfolge fixiert und dann nur die auftretenden Zahlen (samt Vorzeichen) notiert. Man erhält dann für ein solches lineares Gleichungssystem die folgende *erweiterte Matrix*

$$\left(\begin{array}{cccc|c} a_{11} & a_{12} & \dots & a_{1n} & b_1 \\ a_{21} & a_{22} & \dots & a_{2n} & b_2 \\ \vdots & \vdots & & \vdots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} & b_m \end{array} \right)$$

Das Schema der a_{ij} nennt man die *Koeffizientenmatrix*, während der Vektor der b 's als die *rechte Seite* des Gleichungssystems bezeichnet wird. Die Zahl m der Gleichungen finden wir nun als Zahl der *Zeilen* und die Zahl n der Unbekannten als Zahl der *Spalten der Koeffizientenmatrix* wieder.

c. Gauß'sches Eliminationsverfahren. Lineare Gleichungssysteme löst man mit Hilfe des *Gauß'schen Eliminationsverfahrens*: Zulässige Umformungsschritte sind die folgenden elementaren Zeilenumformungen (die mit den kompletten Zeilen der erweiterten Matrix durchgeführt werden):

1. Zeilentausch
2. Multiplikation einer Zeile mit einer Zahl $\neq 0$
3. Addition eines beliebigen reellen Vielfachen einer Zeile zu einer anderen
4. Spaltentausch.

Diese Umformungen bedeuten für das entsprechende Gleichungssystem *Äquivalenzumformungen*, ändern also die Lösungsmenge nicht. (Lediglich beim Spaltentausch muss man die Zuordnung von Spalten und Unbestimmten beachten.)

Begründung: Erstens überführen die obigen Umformungen ein wahres Gleichungssystem wieder in ein wahres, und zweitens können die obigen Umformungen durch gleichartige Umformungen rückgängig gemacht werden: 1. und 4. durch den umgekehrten Tausch, 2. durch Division durch dieselbe Zahl (möglich, da diese $\neq 0$ sein sollte), und 3. durch Subtraktion desselben Vielfachen der unverändert gebliebenen (!) Zeile von der anderen.

d. Der Rang. Man kann mit dem Gauß-Verfahren jedes lineare Gleichungssystem in ein äquivalentes überführen, dessen erweiterte Matrix folgende spezielle Gestalt (die sogenannte Dreiecksform) hat: (Dabei steht * für irgendwelche Einträge.)

$$\left(\begin{array}{cccc|c} a'_{11} & & & & b'_1 \\ 0 & a'_{22} & & * & b'_2 \\ \vdots & & \ddots & & \vdots \\ 0 & 0 & \dots & a'_{rr} & b'_r \\ 0 & 0 & \dots & 0 & 0 \dots 0 & b'_{r+1} \\ \vdots & & & & \vdots & \vdots \\ 0 & 0 & \dots & 0 & 0 \dots 0 & b'_m \end{array} \right)$$

1. Die so entstandene Matrix enthält unter der *Hauptdiagonalen* nur Nullen,
2. die ersten r ($0 \leq r \leq n, m$) Hauptdiagonalelemente a'_{11}, \dots, a'_{rr} sind alle $\neq 0$, und
3. die weiteren Zeilen der Koeffizientenmatrix (ab Nummer $r+1$ bis m) sind (falls überhaupt vorhanden) Nullzeilen.

Die hierbei auftretende Zahl r nennen wir den *Rang* der Koeffizientenmatrix (oder auch des Gleichungssystems). Es gilt $r \leq n$ und $r \leq m$. (Ohne Beweis vermerken wir: Der Rang ist unabhängig von den gewählten Umformungsschritten.)

e. Lösbarkeit. Die $m - r$ Nullzeilen am Ende stellen Gleichungen der Form

$$0 = b'_{r+1}, \dots, 0 = b'_m$$

dar. Diese entscheiden über die Lösbarkeit des gesamten Gleichungssystems:

1. Ist *eine* dieser Gleichungen *nicht erfüllt*, so ist das gesamte Gleichungssystem *unlösbar*.
2. Sind *alle* diese Gleichungen *erfüllt*, so kann man die übrigen r Gleichungen sukzessive nach x_r, \dots, x_1 auflösen (wegen $a'_{rr} \neq 0, \dots, a'_{11} \neq 0$) und so eine Lösung für das Gesamtsystem finden: Das Gleichungssystem ist *lösbar*.

f. Lösungsmenge. Im Falle der Lösbarkeit(!) sind die letzten Gleichungen stets wahr und damit überflüssig. Das Gleichungssystem reduziert sich somit auf die ersten r Gleichungen mit folgender erweiterter Matrix:

$$\left(\begin{array}{cccc|c} a'_{11} & & & & b'_1 \\ 0 & a'_{22} & & * & b'_2 \\ \vdots & & \ddots & & \vdots \\ 0 & 0 & \cdots & a'_{rr} * \dots * & b'_r \end{array} \right)$$

Da hierin die Hauptdiagonalelemente $a'_{ii} \neq 0$ sind, kann man das Gleichungssystem 'von unten nach oben' auflösen und erhält so Formeln für x_r, \dots, x_1 , in denen evtl. die übrigen Unbekannten x_{r+1}, \dots, x_n auftreten. Dies bedeutet, man kann die $n-r$ Variablen x_{r+1}, \dots, x_n *frei wählen*, und dann gemäß den ermittelten Formeln für x_r, \dots, x_1 immer zu einer genau einer Lösung x_1, \dots, x_n vervollständigen. Dies führt zu folgenden fundamentalen Aussagen über die Lösungsgesamtheit, *vorausgesetzt das Gleichungssystem ist lösbar!*

1. Ist $n = r$, so kann man nichts frei wählen und erhält *genau eine* Lösung;
2. ist $n > r$, so kann man $n - r$ Größen frei wählen und erhält so unendlich viele Lösungen. Genauer gilt (vgl. Übungen):

Man erhält eine Parameterdarstellung für die Lösungsmenge, wobei die $n - r$ frei wählbaren Unbekannten x_{r+1}, \dots, x_n die *Parameter* sind und die zugehörigen Vektoren linear unabhängig sind. Damit wird die Lösungsmenge eine Gerade (bei $n - r = 1$), eine Ebene (bei $n - r = 2$) bzw. allgemein eine $n - r$ -dimensionale affine Teilmenge.

Zusammenfassend:

- A) Der Rang r ist höchstens so groß wie die kleinere der beiden Zahlen n und m : $r \leq m$ und $r \leq n$.
- B) $m - r$ ist die Zahl der Lösbarkeitsbedingungen.
 $m = r$ bedeutet: Das Gleichungssystem ist *bei jeder beliebigen rechten Seite* (aber unveränderter Koeffizientenmatrix (a_{ij})) lösbar.
- C) Die Lösungsmenge eines linearen Gleichungssystems ist leer oder eine $n - r$ -dimensionale affine Menge.
 $n = r$ bedeutet: Die Lösungsmenge ist leer oder besteht aus genau einem Punkt.

g. Homogene Gleichungssysteme. Darunter versteht man lineare Gleichungssysteme, bei denen die rechte Seite nur Nullen enthält. Für diese homogenen Systeme gelten die obigen Überlegungen selbstverständlich auch, nur gilt *zusätzlich*:

Ein homogenes System ist immer lösbar.

Denn: $x_1 = \dots = x_n = 0$ ist in jedem Fall eine Lösung, die sog. triviale Lösung. Dadurch werden die obigen Formulierungen im homogenen Falle einfacher. B) entfällt und C) lautet nun:

- C') *Die Lösungsmenge eines homogenen Gleichungssystems ist eine $n - r$ -dimensionale affine Teilmenge, die durch den Koordinatenursprung O verläuft, d. h. ein $n - r$ -dimensionaler Teilvektorraum des \mathbb{R}^n .*

h. Rang und lineare Unabhängigkeit. Sind u_1, \dots, u_n Vektoren im Vektorraum \mathbb{R}^m , deren lineare Unabhängigkeit wir untersuchen wollen, so muss man zeigen, dass die Vektorgleichung $r_1 u_1 + \dots + r_n u_n = o$ *nur* die (triviale) Lösung $r_1 = \dots = r_n = 0$ hat. Schreibt man diese Vektorgleichung $r_1 u_1 + \dots + r_n u_n = o$ einmal vollständig aus, so erhält man für die n Unbekannten r_1, \dots, r_n ein *homogenes* lineares Gleichungssystem, dessen Koeffizientenmatrix

$$A = \left(\begin{array}{c|c|c|c} u_1 & u_2 & \cdots & u_n \end{array} \right)$$

Kombiniert man nun die obigen Ergebnisse, so erhält man folgende Charakterisierungen von Basen (dies sind definitionsgemäß linear unabhängige Erzeugendensysteme):

n Vektoren des \mathbb{R}^m bilden genau dann eine Basis des \mathbb{R}^m , wenn die mit ihnen als Spaltenvektoren gebildete Matrix den Rang $r = n = m$ hat.

Dies bedeutet insbesondere: Basen des \mathbb{R}^m müssen also genau m Vektoren umfassen.

j. Koordinatengleichungen affiner Mengen. Bisher haben wir affine Teilmengen (Geraden, Ebenen, ...) immer durch Parameterdarstellungen

$$\overrightarrow{OX} = a + r_1 u_1 + \dots + r_d u_d$$

beschrieben. Dabei ist a Ortsvektor eines fest gewählten Punktes in der affinen Teilmenge und u_1, \dots, u_d sind d linear unabhängige Richtungsvektoren. Die d Parameter r_1, \dots, r_d sind Variable für beliebige reelle Zahlen. Die Größe d ist die *Dimension* der affinen Teilmenge (Gerade $d = 1$, Ebene $d = 2, \dots$).

Der Vorteil einer Parameterdarstellung ist, dass ihre Form nur von der Dimension d des betrachteten geometrischen Gebildes abhängt, nicht aber von der Dimension des umgebenden Raumes. Außerdem ermöglicht eine Parameterdarstellung eine gute geometrische Veranschaulichung: Man kennt einen Punkt (bzw. seinen Ortsvektor a) sowie ein System unabhängiger Richtungsvektoren u_1, \dots, u_d .

Man kann daher leicht beliebige viele Punkte zu *konstruieren*, die zu der affinen Menge gehören. Es ist jedoch schwierig, von einem beliebig vorgelegten Punkt zu *entscheiden*, ob er zu der affinen Teilmenge gehört oder nicht.

Ein Beispiel. Gegeben eine Ebene e mit der Parameterdarstellung

$$X \in e \iff \overrightarrow{OX} = a + ru + sv = \begin{pmatrix} 2 \\ -1 \\ 0 \end{pmatrix} + r \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix} + s \begin{pmatrix} 0 \\ 1 \\ 2 \end{pmatrix}$$

und die Punkte $P = (1, 4, 5)$ sowie $Q = (3, -2, 1)$. Liegen diese Punkte in der Ebene oder nicht?

Um dies etwa für P zu entscheiden, muss man untersuchen, ob sein Ortsvektor $p = \overrightarrow{OP}$ eine Darstellung in Form der Parameterdarstellung hat, d. h. ob es Parameterwerte $r, s \in \mathbb{R}$ gibt mit

$$p = a + ru + sv.$$

Dies ist ein lineares Gleichungssystem mit 2 Unbekannten (r, s) und der erweiterten Matrix

$$\left(u \mid v \parallel p - a \right) = \left(\begin{array}{cc|c} 1 & 0 & -1 \\ -2 & 1 & 5 \\ 1 & 2 & 5 \end{array} \right).$$

P gehört nun zur Ebene e genau dann, wenn dieses lineare Gleichungssystem *lösbar* ist. (Eine Lösung braucht dafür nicht bestimmt zu werden.) Man führt also das Gauß-Verfahren durch, bis man erkennen kann, ob die Lösbarkeitsbedingung(en) erfüllt sind. Das Gauß-Verfahren ergibt:

$$\left(\begin{array}{cc|c} 1 & 0 & -1 \\ -2 & 1 & 5 \\ 1 & 2 & 5 \end{array} \right) \leftrightarrow \left(\begin{array}{cc|c} 1 & 0 & -1 \\ 0 & 1 & 3 \\ 0 & 2 & 6 \end{array} \right) \leftrightarrow \left(\begin{array}{cc|c} 1 & 0 & -1 \\ 0 & 1 & 3 \\ 0 & 0 & 0 \end{array} \right)$$

Die Lösbarkeitsbedingung ist erfüllt (in der letzten (Null-)Zeile steht auch auf der rechten Seite eine 0): Der Punkt P gehört zu e .

Will man nun feststellen, ob der Punkt Q zu e gehört, so hat man entsprechend vorzugehen und erneut ein lineares Gleichungssystem auf Lösbarkeit zu untersuchen:

$$\left(\begin{array}{cc|c} 1 & 0 & 1 \\ -2 & 1 & -1 \\ 1 & 2 & 1 \end{array} \right) \leftrightarrow \left(\begin{array}{cc|c} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 2 & 0 \end{array} \right) \leftrightarrow \left(\begin{array}{cc|c} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & -2 \end{array} \right)$$

Das System ist offenbar unlösbar: Q gehört nicht zu e . Man erkennt unschwer, dass man dieselben Umformungen durchgeführt hat und sich dabei aber nur auf der rechten Seite etwas ändert. Dies beruht darauf, dass sich zwar die Gauß-Umformungen auf die gesamten Zeilen der erweiterten Matrix auswirken, aber die Entscheidung, *welche* Umformung man durchführt, hängt nur von der Koeffizientenmatrix ab.

Statt mehrfach dieselben Rechnungen durchzuführen, kann man das Problem allgemein lösen. Man untersucht für einen beliebigen Punkt $P = (x, y, z)$, ob er zu e gehört oder nicht. Man muss also das lineare Gleichungssystem mit der erweiterten Matrix

$$\left(u \mid v \parallel p - a \right) = \left(\begin{array}{cc|c} 1 & 0 & x - 2 \\ -2 & 1 & y + 1 \\ 1 & 2 & z \end{array} \right).$$

auf Lösbarkeit untersuchen. (Achtung: Die Unbekannten sind hier r, s und nicht x, y, z .) Führt man nun die Gauß-Elimination durch, so erhält man:

$$\begin{aligned} \left(\begin{array}{cc|c} 1 & 0 & x - 2 \\ -2 & 1 & y + 1 \\ 1 & 2 & z \end{array} \right) &\leftrightarrow \left(\begin{array}{cc|c} 1 & 0 & x - 2 \\ 0 & 1 & y + 1 + 2(x - 2) \\ 0 & 2 & z - (x - 2) \end{array} \right) \leftrightarrow \\ &\leftrightarrow \left(\begin{array}{cc|c} 1 & 0 & x - 2 \\ 0 & 1 & 2x + y - 3 \\ 0 & 0 & z - x + 2 - 2(2x + y - 3) \end{array} \right) \end{aligned}$$

Dieses Gleichungssystem ist genau dann lösbar, wenn der in der Nullzeile auf der rechten Seite stehende Term $z - x + 2 - 2(2x + y - 3) = -5x - 2y + z + 8$ den Wert 0 hat. Also:

$$(x, y, z) \in e \iff -5x - 2y + z + 8 = 0.$$

Dies ist nun eine für jeden Punkt (x, y, z) einfach zu überprüfende Bedingung. Zum Beispiel für den Punkt $P = (1, 4, 5)$ ist diese Gleichung erfüllt: $-5 \cdot 1 - 2 \cdot 4 + 5 + 8 = 0$; P liegt also in der Ebene. Für $Q = (3, -2, 1)$ hingegen gilt die Gleichung nicht: $-5 \cdot 3 - 2 \cdot (-2) + 1 + 8 = -2 \neq 0$; Q gehört nicht zu e . Es gehören also alle die Punkte X zu der Ebene, deren Koordinaten diese Gleichung erfüllen ('lösen'): Die Ebene e ist damit Lösungsmenge dieser einen (Koordinaten-)Gleichung.

Bei der Lösung linearer Gleichungssysteme bestimmt man für die Lösungsmenge eine Parameterdarstellung. Im obigen Beispiel sind wir umgekehrt vorgegangen: Es war eine Ebene e durch eine Parameterdarstellung gegeben und wir haben e als Lösungsmenge einer linearen Gleichung ($-5x - 2y + z + 8 = 0$) dargestellt. Dies ist allgemein möglich. Jedoch kommt man nicht immer mit *einer* Gleichung aus; man braucht im allgemeinen ein lineares Gleichungssystem.

2. Beispiel: Unter welchen Bedingungen gehört ein Punkt $X = (x, y, z)$ zu der Geraden g durch die Punkte $A = (2, 1, -1)$ und $B = (-1, 1, 0)$?

Man bestimmt zunächst eine Parameterdarstellung für g :

$$x = \begin{pmatrix} 2 \\ 1 \\ -1 \end{pmatrix} + r \begin{pmatrix} -3 \\ 0 \\ 1 \end{pmatrix}$$

und untersucht dann das Gleichungssystem mit nur einer Unbekannten (r) und der erweiterten Matrix

$$\left(u \mid x - a \right) = \left(\begin{array}{c|c} -3 & x - 2 \\ 0 & y - 1 \\ 1 & z + 1 \end{array} \right)$$

auf seine Lösbarkeit. Man erhält durch Gauß-Elimination

$$\left(\begin{array}{c|c} -3 & x - 2 \\ 0 & y - 1 \\ 1 & z + 1 \end{array} \right) \leftrightarrow \left(\begin{array}{c|c} -3 & x - 2 \\ 0 & y - 1 \\ 0 & 3(z + 1) + (x - 2) \end{array} \right).$$

Der Rang der Koeffizientenmatrix ist $r = 1$ und die Anzahl der Lösbarkeitsbedingungen daher $m - r = 3 - 1 = 2$: Die Gleichungen $y - 1 = 0$ und $3(z + 1) + (x - 2) = x + 3z + 1 = 0$ müssen erfüllt sein, damit der Punkt $X = (x, y, z)$ auf der Geraden g liegt. g ist damit Lösungsmenge des linearen Gleichungssystems

$$\begin{aligned} y - 1 &= 0 \\ x + 3z + 1 &= 0. \end{aligned}$$

An diesem zweiten Beispiel erkennt man schon recht deutlich den Zusammenhang zwischen der Dimension d der gegebenen affinen Teilmenge, der Dimension m des umgebenden Raumes, sowie der benötigten Zahl von Gleichungen:

Eine affine Teilmenge der Dimension d im \mathbb{R}^m ist Lösungsmenge eines linearen Gleichungssystem bestehend aus $m - d$ Gleichungen.

Wir wollen dies nun allgemein begründen. Gegeben ist eine Parameterdarstellung $x = a + r_1 u_1 + \dots + r_d u_d$ einer d -dimensionalen affinen Teilmenge im \mathbb{R}^m . Ein Punkt $X = (x_1, \dots, x_m) \in \mathbb{R}^m$ gehört zu diesem Raum, wenn das Gleichungssystem mit der erweiterten Matrix

$$\left(\begin{array}{c|ccc} u_1 & \cdots & u_d & \\ \hline & & & x - a \end{array} \right)$$

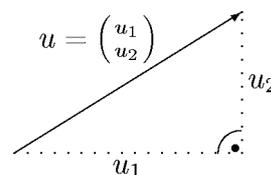
lösbar ist. Dieses System hat die Zeilenzahl m , die Spaltenzahl d und, da u_1, \dots, u_d linear unabhängig sind, den Rang d (siehe h.). Damit erhält man nach der Gauß-Elimination $m - d$ Bedingungen für die Lösbarkeit. Diese Bedingungen sind Gleichungen, in denen die Koordinaten x_i von X nur in erster Potenz vorkommen. Ein Punkt gehört also zu der affinen Teilmenge, wenn seine Koordinaten diese $m - d$ linearen Gleichungen erfüllen, wenn X also eine Lösung dieses Gleichungssystems ist.

§4 Länge, Orthogonalität, Winkel – Skalarprodukt

Wir haben bisher analytische Geometrie allein auf dem Vektorbegriff aufgebaut. Wir haben affine Mengen, ihre gegenseitige Lage, ihre Schnittmengen und Längenverhältnisse betrachtet, nicht jedoch die Längen selbst. Ebenso haben wir den Winkelbegriff in unserem bisherigen Geometrieaufbau nicht behandelt. Dies soll jetzt geschehen. Man nennt diesen Teil der analytischen Geometrie *metrische Geometrie*. Bei der Einführung der wichtigen Grundbegriffe lassen wir uns vom Satz des Pythagoras leiten.

a. Längen. Ist ein Vektor $u = \begin{pmatrix} u_1 \\ u_2 \end{pmatrix}$ in der Ebene gegeben, so berechnet man seine Länge $|u|$ leicht mit Hilfe des Satzes des Pythagoras:

$$|u|^2 = u_1^2 + u_2^2, \quad |u| = \sqrt{u_1^2 + u_2^2}.$$

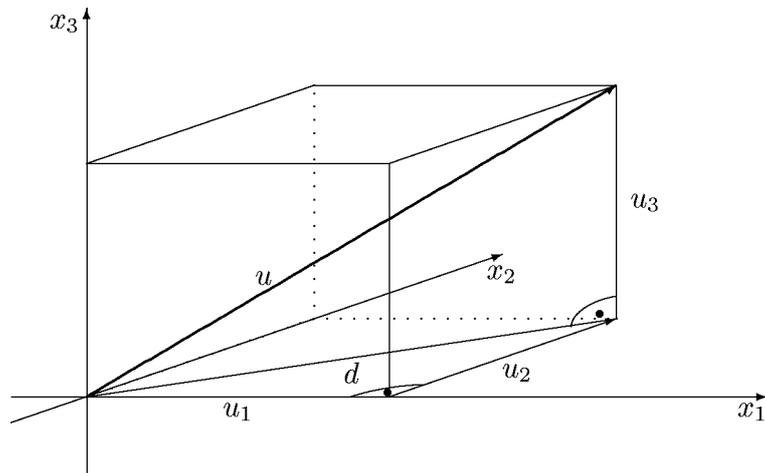


Im dreidimensionalen Fall bestimmt man die Länge eines Vektors $u = \begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix}$ durch zweimalige Anwendung des Satzes des Pythagoras (siehe nachfolgende Skizze). Zunächst berechnet man die Länge d der Diagonale des Bodenrechtecks aus der Beziehung

$$d^2 = u_1^2 + u_2^2$$

und dann damit in dem 'senkrecht' stehenden, ebenfalls rechtwinkligen Dreieck mit Hypotenuse u die Länge von u

$$|u|^2 = d^2 + u_3^2 = u_1^2 + u_2^2 + u_3^2, \quad |u| = \sqrt{u_1^2 + u_2^2 + u_3^2}.$$



Aufgrund dieser Formel liegt die folgende allgemeine Definition nahe:

Definition: Für Vektoren im \mathbb{R}^n definiert man das Längenquadrat und die Länge durch

$$u = \begin{pmatrix} u_1 \\ \vdots \\ u_n \end{pmatrix} \implies |u|^2 = u_1^2 + u_2^2 + \dots + u_n^2, \quad |u| = \sqrt{u_1^2 + \dots + u_n^2},$$

und für Punkte A, B im n -dimensionalen Raum definiert man dann den Abstand $|AB|$ als Länge des Verbindungsvektors

$$|AB| = |\overrightarrow{AB}|$$

Aufgrund dieser Definition erhält man die folgenden Eigenschaften

$$\begin{aligned} |u| = 0 &\iff u = o \\ |ru|^2 &= r^2 |u|^2, \quad |ru| = |r| \cdot |u|. \end{aligned}$$

Bei der ersten Beziehung beachte man, dass eine Summe von (nicht-negativen) Quadratzahlen nur 0 ergeben kann, wenn *alle* Summanden 0 sind. In der zweiten Zeile ist die erste Beziehung unmittelbar nachzurechnen; für die zweite beachte man, dass $\sqrt{r^2} = |r|$ (und nicht $= r$!) ist.

b. Skalarprodukt. Wir werden im Laufe der Zeit sehen, dass dieser Begriff das Fundament der gesamten *metrischen* Geometrie ist. Nicht nur der schon eingeführte Längenbegriff erweist sich als ein Spezialfall des Skalarproduktes, sondern auch der vorerst zurückgestellte Winkelbegriff basiert darauf. *Das Skalarprodukt ist der fundamentale Begriff der metrischen Geometrie. Aus ihm lassen sich Länge und Winkel ableiten.* Ein besonderer Vorzug des Skalarproduktes sind seine einfachen Eigenschaften.

Definition: Für je zwei Vektoren im \mathbb{R}^n definiert man das Skalarprodukt $u \cdot v$ durch

$$u \cdot v = \begin{pmatrix} u_1 \\ \vdots \\ u_n \end{pmatrix} \cdot \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} = u_1 v_1 + u_2 v_2 + \dots + u_n v_n.$$

Spezialisiert man hierin $u = v$, so erkennt unmittelbar die Beziehung zwischen Skalarprodukt und Länge:

$$|u|^2 = u \cdot u :$$

Das Längenquadrat eines Vektors ist das Skalarprodukt des Vektors mit sich selbst. Diese Beschreibung ist deshalb so vorteilhaft, weil das Skalarprodukt sehr einfachen Regeln genügt. Es gilt:

Eigenschaften des Skalarproduktes: Es gelten die folgenden Gesetzmäßigkeiten für beliebige Vektoren u, v, w und beliebige reelle Zahlen r :

$$\begin{aligned} u \cdot v &= v \cdot u, \\ (ru) \cdot v &= r \cdot (u \cdot v), \\ (u + v) \cdot w &= u \cdot w + v \cdot w, \\ u \neq o &\implies u \cdot u > 0. \end{aligned}$$

Alle diese Gesetzmäßigkeiten rechnet man aufgrund der Definition nach. Nach dem ersten Gesetz ist das Skalarprodukt *symmetrisch*. Das zweite Gesetz erlaubt es, in derartigen Termen Klammern wegzulassen. Aufgrund der Distributivität (drittes Gesetz) verhält sich das Skalarprodukt wie man es von einem Produkt erwartet. Man kann daher weitgehend wie gewohnt rechnen. Wegen der Symmetrie gelten die zweite und dritte Regel auch für den zweiten Faktor.

Achtung: Dennoch ist Vorsicht geboten. Da das Skalarprodukt zweier Vektoren *kein* Vektor, sondern ein *Skalar*, also eine reelle Zahl ist, macht ein dreifaches Skalarprodukt keinen Sinn!

$$u \cdot v \cdot w \quad \text{ist sinnlos!}$$

Aus diesen Regeln für das Skalarprodukt kann man Konsequenzen für die Länge von Vektoren ziehen. So kann man die Länge eines Summen- oder Differenzvektors berechnen:

$$\begin{aligned} |u \pm v|^2 &= (u \pm v) \cdot (u \pm v) = u \cdot (u \pm v) \pm v \cdot (u \pm v) \\ &= u \cdot u \pm u \cdot v \pm v \cdot u + v \cdot v \\ &= |u|^2 \pm 2u \cdot v + |v|^2 \end{aligned}$$

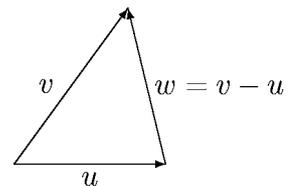
Dies stellt ein Analogon zu den ersten beiden binomischen Formeln dar. Das Analogon der dritten binomischen Formel lautet:

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

Diese Formeln haben vielfältige geometrische Konsequenzen (siehe Übungen (7), Aufgabe 2).

Wir wollen hier einmal die zweite binomische Formel geometrisch ausdeuten: Ist $w = v - u$, so bilden die drei Vektoren die drei Seiten eines Dreiecks (siehe Skizze). Die zweite binomische Formel liefert dann

$$|w|^2 = |v|^2 + |u|^2 - 2u \cdot v.$$



Man erkennt eine Ähnlichkeit mit dem Satz des Pythagoras, nur dass hier bei beliebigen Dreiecken ein *Korrekturterm* auftritt, nämlich $-2u \cdot v$.

Man sieht daran:

In einem Dreieck gilt der Satz des Pythagoras $|w|^2 = |v|^2 + |u|^2$ genau dann, wenn $u \cdot v = 0$ ist.

c. Orthogonalität. Da der Satz des Pythagoras bekanntlich genau für *rechtwinklige* Dreiecke gilt, sieht man hier, dass mit dem Skalarprodukt auch die *Orthogonalität* (Rechtwinkligkeit) beschrieben werden kann. Die folgende Definition ist daher sinnvoll:

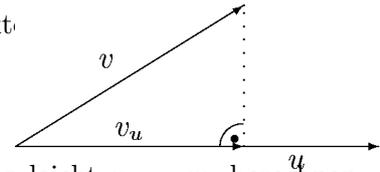
Definition: Zwei Vektoren des \mathbb{R}^n nennt man *orthogonal*, in Zeichen $u \perp v$, wenn ihr Skalarprodukt Null ist:

$$u \perp v \iff u \cdot v = 0.$$

Man beachte, dass bei dieser Definition der Nullvektor o zu *jedem* Vektor orthogonal ist.

Wir wollen nun den engen Zusammenhang des Skalarproduktes mit der vor allem in der Physik so wichtigen *orthogonalen Projektion* eines Vektors v auf einen anderen $u \neq 0$ aufzeigen. Dabei ist der Projektionsvektor (siehe Skizze) charakterisiert durch:

$$v_u = ru \text{ ist Vielfaches von } u \quad \text{und} \quad v - v_u \perp u.$$



Mit den Eigenschaften des Skalarproduktes kann man daraus sehr leicht $v_u = ru$ berechnen, denn:

$$(v - ru) \cdot u = 0 \iff v \cdot u - ru \cdot u = 0 \iff v \cdot u = r \cdot |u|^2.$$

Diese Gleichung lässt sich (wegen $|u|^2 \neq 0$) nach r auflösen und ergibt die folgende allgemeine Formel für r und damit v_u :

$$v_u = ru \quad \text{mit} \quad r = \frac{v \cdot u}{u \cdot u} = \frac{v \cdot u}{|u|^2}.$$

In mehr geometrischer Sprechweise ist dies die Bestimmung eines *Lotfußpunktes*: Man betrachtet die Richtung von u als den ‘Boden’ und lässt von der Spitze von v ein *Lot* herunterhängen: Die Spitze von v_u ist dann der Lotfußpunkt. (Zur Berechnung von Lotfußpunkten und ihrer Bedeutung siehe Übungen (6).)

Man kann die obige Bestimmung des Projektionsvektors v_u auch benutzen, um eine geometrische Deutung des Skalarproduktes allgemein (nicht nur für $u \cdot u$ und für $u \cdot v = 0$) zu geben. Dazu gehen wir von der obigen Bestimmung von v_u aus

$$v_u = \frac{v \cdot u}{u \cdot u} \cdot u$$

und bilden das Skalarprodukt mit u :

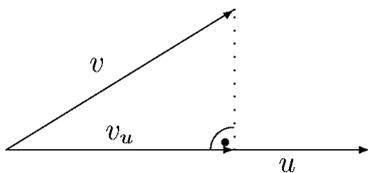
$$v_u \cdot u = \frac{v \cdot u}{u \cdot u} \cdot u \cdot u = v \cdot u.$$

Dies bedeutet zunächst, dass das Skalarprodukt $v \cdot u$ übereinstimmt mit dem Skalarprodukt $v_u \cdot u$ der beiden Vektoren v_u und u . Diese sind nach Ansatz linear abhängig: $v_u = ru$. Und für linear abhängige Vektoren ist das Skalarprodukt bis auf das Vorzeichen gleich dem Produkt der Längen der Vektoren:

$$(ru) \cdot u = r \cdot u \cdot u = r \cdot |u|^2 = (r|u|) \cdot |u| = \begin{cases} +|ru| \cdot |u| & \text{falls } r \geq 0, \\ -|ru| \cdot |u| & \text{falls } r < 0. \end{cases}$$

Dabei bedeutet $r > 0$ gerade, dass $v_u = ru$ und u gleiche *Orientierung* haben. Also

$$v \cdot u = v_u \cdot u = \begin{cases} +|v_u| \cdot |u| & \text{falls } u, v_u \text{ gleich orientiert,} \\ -|v_u| \cdot |u| & \text{falls } u, v_u \text{ entgegengesetzt orientiert.} \end{cases} \quad (*)$$



d. Winkel. Die letzten beiden Skizzen zeigen deutlich einen Zusammenhang zwischen dem Vorzeichen von $v \cdot u$ und der Art des Winkels α zwischen u und v : Ist der Winkel spitz, so ist das

Skalarprodukt positiv, andernfalls negativ. Tatsächlich kann aus dem Wert des Skalarprodukts sogar der Winkel genau bestimmt werden. Wir gehen von zwei Vektoren $u, v \neq o$ und der letzten Beschreibung (*) des Skalarproduktes aus: $v \cdot u = \pm |v_u| \cdot |u|$. Nun bestimmen die Vektoren v_u und v ein rechtwinkliges Dreieck, so dass bei spitzem Winkel α zwischen u und v (siehe vorangehende linke Skizze) gemäß der Definition des Cosinus gilt:

$$\cos(\alpha) = \frac{|v_u|}{|v|} \stackrel{(*)}{=} \frac{v \cdot u}{|u| |v|}.$$

Im Falle eines stumpfen Winkels α (rechte Skizze) bilden die Vektoren v_u und v ebenfalls ein rechtwinkliges Dreieck, aber mit dem Winkel $180^\circ - \alpha$. In diesem Falle erhält man wiederum aus (*)

$$\cos(180^\circ - \alpha) = \frac{|v_u|}{|v|} \stackrel{(*)}{=} -\frac{v \cdot u}{|u| \cdot |v|}.$$

Da aber $\cos(180^\circ - \alpha) = -\cos(\alpha)$ ist (siehe Definition des Cosinus am Einheitskreis), erhält man auch in diesem Falle die fundamentale Beziehung für den Winkel α zwischen zwei Vektoren $u, v \neq o$:

$$\cos(\alpha) = \frac{u \cdot v}{|u| \cdot |v|}, \quad \alpha = \arccos\left(\frac{u \cdot v}{|u| \cdot |v|}\right)$$

Zugleich gibt diese Beziehung eine weitere geometrische Beschreibung des Skalarproduktes für Vektoren $u, v \neq o$:

$$u \cdot v = |u| \cdot |v| \cdot \cos(\alpha) \quad \text{mit dem Winkel } \alpha \text{ zwischen } u \neq o \text{ und } v \neq o.$$

Für $u = o$ oder $v = o$ gibt es keinen Winkel α zwischen u, v , aber dann gilt bekanntlich $u \cdot v = 0$.

e. Normalenvektoren und Koordinatengleichungen. In Abschnitt j. haben wir die Beschreibung affiner Mengen durch Koordinatengleichungen behandelt. Während Parameterdarstellungen affiner Mengen schon vom Ansatz her geometrische Bedeutung haben, ist die geometrische Bedeutung einer Koordinatengleichung $a_1 x_1 + \dots + a_n x_n = b$ unklar. Diese wollen wir nun herausarbeiten.

Sei also die affine Menge M die Lösungsmenge einer linearen Gleichung

$$X = (x_1, \dots, x_n) \in M \iff a_1 x_1 + \dots + a_n x_n = b,$$

wobei nicht alle $a_i = 0$ seien. Damit ist M dann eine $n - 1$ -dimensional affine Menge im \mathbb{R}^n . Im Falle $n = 3$ ist M also eine Ebene. Fasst man die Koeffizienten der linearen Gleichung zu

einem Vektor $a = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}$ zusammen und bezeichnet man den Ortsvektor von X mit $x: x =$

$\overrightarrow{OX} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$, so kann man die lineare Gleichung mit Hilfe des Skalarproduktes folgendermaßen schreiben:

$$X = (x_1, \dots, x_n) \in M \iff a \cdot x = b.$$

M besteht also aus allen Punkten X , deren Ortsvektor x mit einem festen Vektor a ein festes Skalarprodukt b hat. Welche geometrische Bedeutung hat nun dieser Vektor a ?

Wir betrachten dazu beliebige Punkte $P, Q \in M$. Dann gilt für deren Ortsvektoren p, q

$$a \cdot p = b = a \cdot q, \text{ also } a \cdot (q - p) = 0.$$

Dies bedeutet

$$a \cdot \overrightarrow{PQ} = 0 \text{ für alle } P, Q \in M,$$

in Worten: a ist senkrecht zu jedem Vektor \overrightarrow{PQ} , $P, Q \in M$, d. h. zu ganz M :

$$M = \{(x_1, \dots, x_n) \in \mathbb{R}^n \mid a_1x_1 + \dots + a_nx_n = b\} \implies \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} \perp M.$$

Man nennt einen Vektor, der $\neq o$ ist und zu einer affinen Menge M orthogonal ist, einen *Normalenvektor* von M .

Man beachte, dass diese Überlegungen auch gelten, wenn M die Lösungsmenge eines Gleichungssystems aus mehreren Gleichungen ist; in diesem Falle ist der Koeffizientenvektor *jeder* der Gleichungen ein Normalenvektor von M .

Der Hauptanwendungsbereich ist aber der Fall einer affinen Menge M der Dimension $n-1$, die durch *eine* Gleichung beschrieben werden kann, und hier insbesondere die Bestimmung von Lotfußpunkten und Abständen.

f. Abstände. Unter dem *Abstand* $d(P, Q)$ zweier Punkte P, Q verstehen wir die Länge des Verbindungsvektor \overrightarrow{PQ} :

$$d(P, Q) = |\overrightarrow{PQ}|.$$

Wir wollen nun den Abstand eines Punktes P von affinen Mengen M (Geraden, Ebenen) untersuchen. Darunter versteht man den *kürzesten* Abstand zwischen P und den Punkten der affinen Menge. Dieser kürzeste Abstand ist die Länge des Lotes von P auf die affine Menge M . (Nach dem Satz des Pythagoras ist der Abstand zwischen P und jedem anderen Punkt der affinen Menge größer. Begründen Sie dies!)

Der Abstand eines Punktes P von einer affinen Menge M ist die Länge des Lotes von P auf M .

Man muss also zur Abstandsbestimmung das Lot von einem Punkt P auf eine affine Menge M bestimmen. Dazu sind nach unserem bisherigen Vorgehen folgende Schritte nötig:

1. Bestimmung einer Parameterdarstellung der affinen Menge:

$$X \in M \iff \overrightarrow{OX} = a + \sum_{i=1}^d r_i u_i \quad \text{für geeignete } r_i \in \mathbb{R}.$$

2. Lotfußpunktbedingung: $\overrightarrow{PX} \perp M$, d. h.

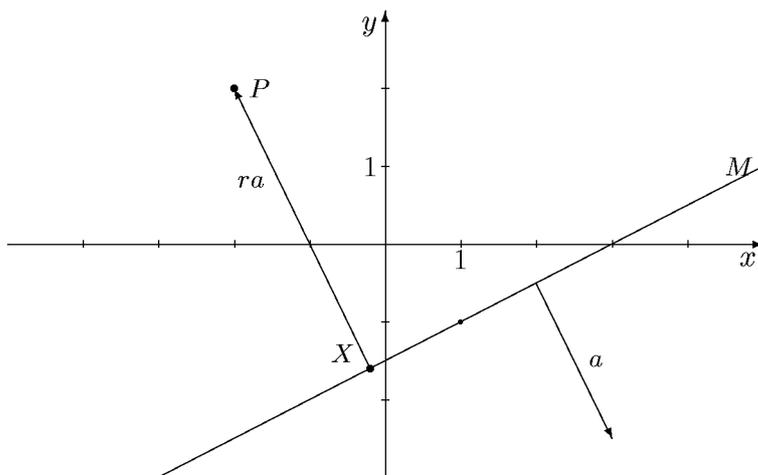
$$(a - p + \sum_{i=1}^d r_i u_i) \cdot u_j = 0 \quad (j = 1, \dots, d).$$

Man erhält so ein lineares Gleichungssystem mit d Gleichungen für d Unbekannte, bei einer Ebene M ($d = 2$) also ein 2×2 -Gleichungssystem.

Wir wollen nun eine andere Methode kennenlernen, die angewendet werden kann, wenn eine $n-1$ -dimensionale affine Menge $M \subset \mathbb{R}^n$ als Lösungsmenge *einer* linearen Gleichung gegeben ist: $a_1x_1 + \dots + a_nx_n = b$, und damit ein Normalenvektor $a \neq o$ für M bekannt ist. Der Vektor \overrightarrow{XP} vom Lotfußpunkt $X \in M$ zu P ist definitionsgemäß orthogonal zu M , also ist \overrightarrow{XP} ein Normalenvektor von M . Da durch die Gleichung für M bereits ein Normalenvektor a bekannt ist, muss gelten

$$X \text{ Lotfußpunkt von } P \text{ auf } M \implies \overrightarrow{XP} = r \cdot a.$$

(Hierbei ist wesentlich, dass M die Dimension $n - 1$ hat und daher alle Normalenvektoren Vielfache voneinander sind.) Gesucht ist nun der eine reelle Parameter r .



Zunächst gilt

$$\overrightarrow{XP} = r \cdot a \iff p - x = r \cdot a \iff x = p - r \cdot a.$$

Da X als Lotfußpunkt in M liegt, erfüllt sein Ortsvektor x die Gleichung für M , d. h. es gilt $a \cdot x = b$. Setzt man hierin die vorangehende Darstellung für x ein, so erhält man *eine* Gleichung für die *eine* Unbekannte r :

$$\begin{aligned} b &= a \cdot x = a \cdot (p - r \cdot a) = a \cdot p - r \cdot a \cdot a \\ \iff r \cdot |a|^2 &= a \cdot p - b \iff r = \frac{a \cdot p - b}{|a|^2} \quad (|a| \neq 0!). \end{aligned}$$

Damit sind r und der Lotfußpunkt explizit berechenbar. Und dann natürlich auch der Abstand $d(P, M)$ des Punktes P von der affinen Menge M :

$$d(P, M) = |\overrightarrow{XP}| = |r| \cdot |a| = \frac{|a \cdot p - b|}{|a|^2} \cdot |a| = \frac{|a \cdot p - b|}{|a|}.$$

Man beachte die Bedeutung des Terms $a \cdot p - b$ im Zähler der obigen Abstandsformel:

$$a \cdot p - b = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} \cdot \begin{pmatrix} p_1 \\ \vdots \\ p_n \end{pmatrix} - b = a_1 p_1 + \dots + a_n p_n - b,$$

m. a. W. diesen Term $a \cdot p - b$ erhält man, wenn man die Koordinaten des Punktes P in die linke Seite der Gleichung $a_1 x_1 + \dots + a_n x_n - b = 0$ für M einsetzt. Liegt der Punkt P in M , erfüllt er die Gleichung und es ergibt sich als Zähler 0; der Abstand von P zu M ist dann natürlich 0. Allgemein ergibt sich ein anderer Wert; dividiert man diesen durch die Länge $|a|$ von a , so erhält man den Abstand, allerdings mit einem zusätzlichen Vorzeichen. Dieses Vorzeichen ist auch das Vorzeichen des oben bestimmten Parameters r und gibt damit an, auf welcher Seite der affinen Menge M der Punkt P liegt: Ist $a \cdot p - b > 0$, also $r > 0$, so liegt der Punkt P auf der Seite von M , zu der der Normalenvektor a weist, ist dagegen $a \cdot p - b < 0$, so liegt P auf der anderen Seite, und ist $a \cdot p - b = 0$, so liegt P in M . Man sagt daher,

$\frac{a \cdot p - b}{ a }$ ist der <i>orientierte</i> Abstand des Punktes P von M .
--

Sein Betrag ist der gewöhnliche Abstand und sein Vorzeichen gibt – wie oben erläutert – an, auf welcher Seite von M der Punkt P liegt.

Beispiel: Wir wollen dies einmal beispielhaft an der skizzierten ebenen Situation durchführen: $M = g$ sei die skizzierte Gerade im \mathbb{R}^2 . Wir zeigen dabei zugleich den Nutzen der Normalenvektoren bei der Erstellung einer Gleichung.

1. Aus den beiden Punkten $(+1, -1)$ und $(3, 0)$ auf der Geraden g entnimmt man als Richtungsvektor der Geraden $u = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$.

2. Ein Normalenvektor ist daher $a = \begin{pmatrix} 1 \\ -2 \end{pmatrix}$.

3. Eine Gleichung für g ist daher $1 \cdot x - 2 \cdot y = b$. b bestimmt man, indem man einen Geradenpunkt (etwa $(3, 0)$) einsetzt: $3 - 2 \cdot 0 = b$, also $b = 3$: $x - 2y - 3 = 0$ ist eine Gleichung für die Gerade g .

4. Der Abstand des Punktes $P = (-2, 2)$ von g ist daher

$$d(P, g) = \frac{|-2 - 2 \cdot 2 - 3|}{\sqrt{1^2 + 2^2}} = \frac{9}{\sqrt{5}} \approx 4,025.$$

5. Der Wert r ist

$$r = \frac{-2 - 2 \cdot 2 - 3}{1^2 + 2^2} = -\frac{9}{5} = -1,8.$$

r ist negativ, entsprechend der Tatsache, dass der Normalenvektor *nicht* zu der Seite der Geraden weist, auf der der Punkt P liegt.

6. Mit der Kenntnis von r kann man den Lotfußpunkt bestimmen:

$$x = p - ra = \begin{pmatrix} -2 \\ 2 \end{pmatrix} + \frac{9}{5} \begin{pmatrix} 1 \\ -2 \end{pmatrix} = \begin{pmatrix} -1/5 \\ -8/5 \end{pmatrix}.$$