

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

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

Unterrichtsbegleitende Skripten sowie
Übungen mit ausführlichen Lösungen

Norbert Klingen

Köln 1998

Inhalt

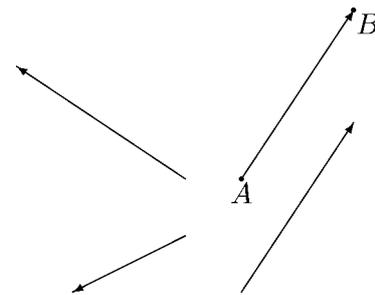
| | |
|---|-----------|
| §1 Von der Anschauung zum Vektorraumbegriff | 1 |
| a. Punkte, Pfeile und Vektoren..... | 1 |
| b. Koordinaten..... | 1 |
| c. Vektorsumme und skalare Multiplikation..... | 2 |
| d. Der Vektorraumbegriff..... | 4 |
| §2 Punkte, Geraden und Ebenen | 5 |
| a. Punkte und Ortsvektoren..... | 5 |
| b. Geraden..... | 6 |
| c. Ebenen..... | 7 |
| d. Affine Koordinatensysteme..... | 8 |
| §3 Lineare Gleichungssysteme und das Gauß'sche Eliminationsverfahren | 8 |
| a. Matrixschreibweise..... | 8 |
| b. Gauß-Elimination und der Rang..... | 9 |
| c. Lösbarkeit und Lösungsvielfalt..... | 9 |
| d. Spezialfälle..... | 10 |
| §4 Gleichungsdarstellung affiner Teilmengen | 11 |
| a. Parameterdarstellungen..... | 11 |
| b. Ein Beispiel..... | 11 |
| c. Gleichungsdarstellung affiner Teilmengen..... | 12 |
| §5 Lineare Unabhängigkeit | 13 |
| a. Parameterdarstellungen und lineare Unabhängigkeit..... | 13 |
| b. Koeffizientenvergleich..... | 13 |
| c. Geometrische Anwendungen..... | 14 |
| d. Rang und lineare Unabhängigkeit..... | 15 |
| §6 Länge, Skalarprodukt, Orthogonalität | 16 |
| a. Längen..... | 16 |
| b. Skalarprodukt..... | 17 |
| c. Orthogonalität..... | 18 |
| d. Normalenvektoren und Gleichungen..... | 18 |
| e. Abstandsberechnungen..... | 19 |

§1 Von der Anschauung zum Vektorraumbegriff

a. Punkte, 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* (siehe Skizze). Jedoch muß man deutlich zwischen Pfeil und Vektor unterscheiden.

Einen Pfeil kann man am einfachsten durch seinen Anfangspunkt und seinen Endpunkt (in der Skizze A und B) festlegen:

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



Dadurch wird der Pfeil in *Länge*, *Richtung* und *Orientierung* (Anfang/Ende) vollständig festgelegt. Darüberhinaus ist aber auch seine konkrete *Lage* im Raum fixiert.

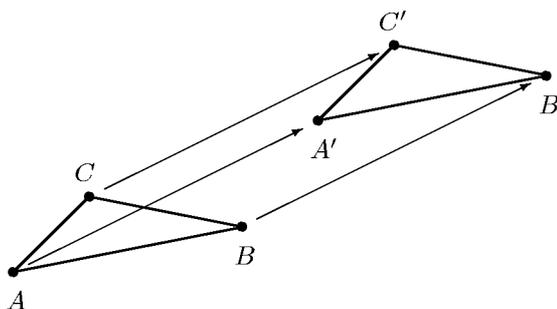
Nun ist aber aus der Physik bekannt, daß es bei den dort auftretenden *Vektoren* gerade nicht auf die Lage des Pfeils ankommt, sondern nur auf *Länge*, *Richtung* und *Orientierung*. Man kommt also vom Pfeil zum Vektor, indem man von der Lage des Pfeils abstrahiert: Unabhängig von ihrer Lage bestimmen verschiedene Pfeile denselben Vektor, wenn sie in Länge, Richtung und Orientierung übereinstimmen:

Zwei Pfeile heißen *vektorgleich*, wenn sie in Länge, Richtung und Orientierung übereinstimmen. Vektorgleiche Pfeile stellen *denselben Vektor* dar.

Sind A und A' zwei Punkte, so bezeichnen wir mit dem Symbol $\overrightarrow{AA'}$ den *Vektor*, der durch den Pfeil von A nach A' bestimmt wird.

Innerhalb der Mathematik kann man den Unterschied zwischen Pfeil und Vektor an den *Verschiebungen* verdeutlichen. Verschiebt man etwa ein Dreieck ABC in ein Dreieck $A'B'C'$, so erhält man Pfeile wie in nebenstehender Skizze. Diese drei verschiedenen Pfeile sind aber vektorgleich, sie bestimmen denselben Vektor, weil sie zu derselben Verschiebung gehören. Man kann also Vektoren als Verschiebungen verstehen. Sehr vereinfacht formuliert:

Vektoren = Verschiebungen

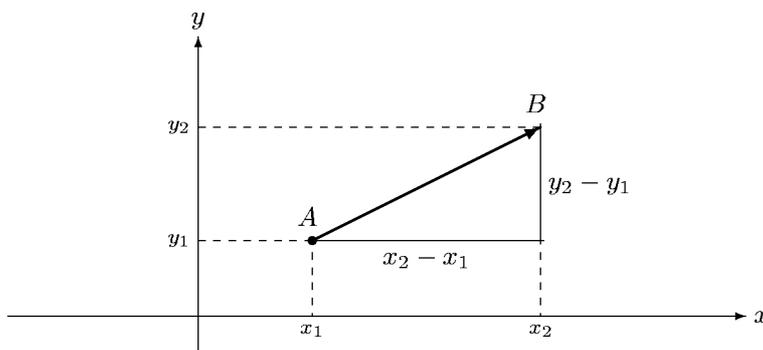


In der obigen Skizze bestimmen alle drei eingezeichneten Pfeile denselben Vektor, den zugehörigen *Verschiebungsvektor*:

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

b. Koordinaten. Wir wollen nun Punkte und Vektoren nicht geometrisch, sondern rechnerisch behandeln. Diese Methode kennzeichnet die *analytische Geometrie*, die Gegenstand dieses Semesters ist. Wir benötigen dazu ein Koordinatensystem, wie wir es aus der Einführungsphase kennen. Darauf aufbauend haben wir dann Punkte durch ihre Koordinaten beschrieben, etwa $A = (x_1, y_1)$, $B = (x_2, y_2)$. Soweit ist dies wohlbekannt.

Wie aber kann man nun den Vektor \overrightarrow{AB} rechnerisch erfassen? Nach Definition ist der Vektor durch Länge, Richtung und Orientierung festgelegt. Aber diese, insbesondere die Richtung, sind rechnerisch nicht so leicht zu behandeln wie die folgende Überlegung: Man ergänzt einen Pfeil, der in einem Koordi-

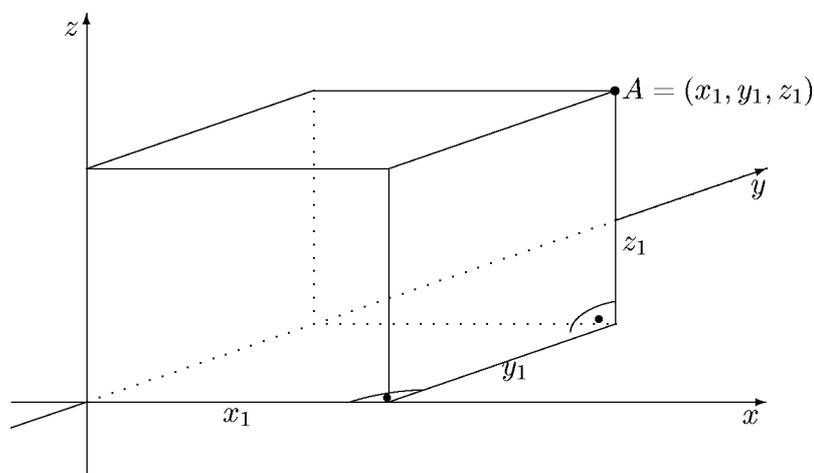


natensystem gegeben ist, durch achsenparallele Katheten zu einem rechtwinkligen Dreieck (siehe Skizze). Länge und Richtung des Vektors, der die Hypotenuse des rechtwinkligen Dreiecks darstellt, sind durch die Katheten festgelegt. Dies ist ähnlich wie bei der Definition des Anstiegs einer Geraden, aber anders als dort kommt es nicht nur auf das Verhältnis der Kathetenlängen zueinander, sondern auf *jede* der beiden einzelnen Katheten an, da die Länge der Hypotenuse festgelegt sein soll. Überdies ist wegen der festgelegten Orientierung die Reihenfolge der Punkte wichtig, so daß auch in den Termen $y_2 - y_1$ bzw. $x_2 - x_1$ die Reihenfolge der Indizes entscheidend ist.

Sind zwei Punkte $A = (x_1, y_1)$ und $B = (x_2, y_2)$ in einem Koordinatenkreuz gegeben, so ist der Vektor \overrightarrow{AB} eindeutig festgelegt durch das Zahlenpaar $\begin{pmatrix} x_2 - x_1 \\ y_2 - y_1 \end{pmatrix}$. Man nennt dies die *Koordinaten des Vektors*. Sie werden spaltenweise geschrieben, um sie von Punktkoordinaten zu unterscheiden.

$$A = (x_1, y_1), B = (x_2, y_2) \implies \overrightarrow{AB} = \begin{pmatrix} x_2 - x_1 \\ y_2 - y_1 \end{pmatrix}$$

All diese Überlegungen kann man auch anwenden, um den uns umgebenden *Raum* zu erfassen. Dazu benötigt man dann ein Koordinatensystem mit *drei* Koordinatenachsen, der x - und y -Achse wie bekannt und eine dritte Achse, die z -Achse, die zu den ersten beiden senkrecht verläuft.



Punkte werden dann durch *drei* Koordinaten beschrieben: $A = (x_1, y_1, z_1)$, und für Vektoren \overrightarrow{AB} ergibt sich dann entsprechend

$$A = (x_1, y_1, z_1), B = (x_2, y_2, z_2) \implies \overrightarrow{AB} = \begin{pmatrix} x_2 - x_1 \\ y_2 - y_1 \\ z_2 - z_1 \end{pmatrix}$$

Für einen beliebigen Punkt $A = (x_1, y_1, z_1)$ führt man den sog. *Ortsvektor* ein als den Verbindungsvektor vom Koordinatenursprung $O = (0, 0, 0)$ zum Punkt A . Wir erhalten:

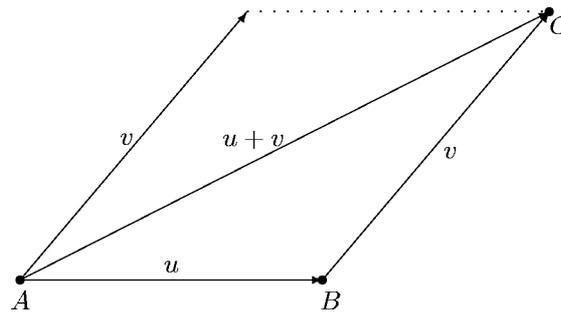
$$\text{Der Ortsvektor von } A: a = \overrightarrow{OA} = \begin{pmatrix} x_1 \\ y_1 \\ z_1 \end{pmatrix}.$$

Man kann also die Punkte in der Ebene oder im Raum durch ihre Ortsvektoren beschreiben, und umgekehrt kann man jeden Vektor als Ortsvektor eines eindeutig bestimmten Punktes darstellen.

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: $u = \overrightarrow{AB}$, $v = \overrightarrow{BC} \Rightarrow u + v = \overrightarrow{AC}$.



Interpretiert man Vektoren als Verschiebungen (der Ebene oder des Raumes), so kann man die Vektorsumme besonders einfach beschreiben. Bei einer Verschiebung gemäß Vektor u kommt man von A zum Punkt B ; verschiebt *danach* gemäß 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:

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

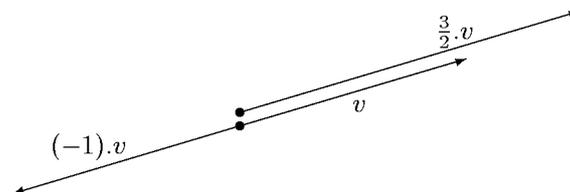
Man erhält also die x -, y - bzw. z -Koordinate des Summenvektors $u + v$ als Summe der entsprechenden Koordinaten von u und v .

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

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

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

Anschaulich bedeutet dies bei $r > 0$ eine Änderung der Länge des Vektors um den Faktor r *ohne Änderung von Richtung oder Orientierung*. Eine Multiplikation mit -1 bedeutet lediglich eine *Umkehr der Orientierung* ohne Längen- oder Richtungsänderung.



Statt $(-1) \cdot v$ schreibt man auch einfach $-v$. Dies ist der sog. *Gegenvektor* von v ; er unterscheidet sich von v nur durch die Umkehr der Orientierung. Ist v durch zwei Punkte A, B gegeben, so lässt sich der Gegenvektor besonders einfach beschreiben; man vertauscht einfach Anfangs- und Endpunkt:

$$v = \overrightarrow{AB} \Rightarrow -v = \overrightarrow{BA}$$

Addiert man nun einen Vektor und seinen Gegenvektor, so erhält man einen Vektor mit gleichem Anfangs- und Endpunkt, den *Nullvektor*:

$$v + (-v) = \overrightarrow{AB} + \overrightarrow{BA} = \overrightarrow{AA} = o, \quad (-v) + v = \overrightarrow{BA} + \overrightarrow{AB} = \overrightarrow{BB} = o.$$

Der Nullvektor wird durch *jeden* Pfeil mit gleichem Anfangs- und Endpunkt dargestellt:

$$o = \overrightarrow{AA} = \overrightarrow{BB}$$

Vorsicht: Der Nullvektor ist kein Punkt! Es gibt (unendlich) viele Punkte, aber nur einen Nullvektor! Man beachte auch, daß der Nullvektor keine Richtung hat.

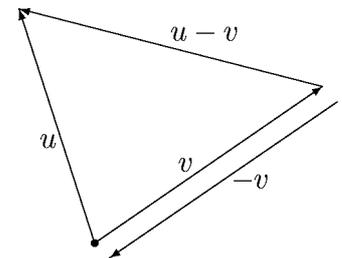
Dieser Nullvektor spielt bei Vektoren dieselbe Rolle, wie die Null 0 bei den Zahlen:

$$o + v = \overrightarrow{AA} + \overrightarrow{AB} = \overrightarrow{AB} = v, \quad v + o = \overrightarrow{AB} + \overrightarrow{BB} = \overrightarrow{AB} = v.$$

Mit Hilfe des Gegenvektors kann man bei Vektoren die Subtraktion genauso einführen wie bei Zahlen:

$$u - v = u + (-v).$$

Geometrisch erhält man den Differenzvektor $u - v$ wie nebenstehend skizziert. Daraus ergibt sich auch die folgende Beziehung zwischen einem Vektor und den Ortsvektoren von Anfangs- und Endpunkt:



$$\overrightarrow{AB} = \overrightarrow{OB} - \overrightarrow{OA}$$

d. Der Vektorraumbegriff. Die obigen Analogien zwischen Vektoren und Zahlen beruhen darauf, daß für die Vektorsumme und skalare Multiplikation eine Reihe von Gesetzmäßigkeiten gelten, wie sie uns von den Zahlen her vertraut sind.

Um diese Eigenschaften 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 Gesetzmäßigkeiten:

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 $u \in V$ gibt es ein $-u \in V$ (genannt 'Gegenvektor' von u) mit der Eigenschaft:

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

Und nun für die skalare Multiplikation:

(SM) (Skalare Multiplikation)

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

(A) (Assoziativgesetz)

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

(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)

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

Der Vorteil der Vektoren gegenüber den Punkten ist die Gültigkeit dieser Gesetze, die uns aus dem Bereich der *Zahlen* wohlvertraut sind. Man kann also über weite Strecken mit Vektoren genauso *rechnen* wie mit Zahlen. Es gibt aber natürliche auch gravierende Unterschiede, insbesondere:

Es gibt keine Division von Vektoren!

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 ' $u + v$ ' und eine skalare Multiplikation ' $r \cdot u$ ' definiert ist, so daß alle oben genannten Eigenschaften erfüllt sind.

Aufgrund unserer einführenden Überlegungen kennen wir eine Reihe von Vektorräumen: Der 2- und 3-dimensionale Anschauungsraum, der Raum der Verschiebungen, die Menge der Tripel reeller Zahlen

$$\mathbb{R}^3 = \left\{ \begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix} \mid u_1, u_2, u_3 \in \mathbb{R} \right\}$$

wie sie als Koordinaten auftraten. Dieses letzte Beispiel kann man nun unmittelbar verallgemeinern zu den Vektorräumen

$$\mathbb{R}^n = \left\{ \begin{pmatrix} u_1 \\ \vdots \\ u_n \end{pmatrix} \mid u_1, \dots, u_n \in \mathbb{R} \right\}$$

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:

$$\begin{pmatrix} u_1 \\ \vdots \\ u_n \end{pmatrix} + \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} = \begin{pmatrix} u_1 + v_1 \\ \vdots \\ u_n + v_n \end{pmatrix}, \quad r \cdot \begin{pmatrix} u_1 \\ \vdots \\ u_n \end{pmatrix} = \begin{pmatrix} r u_1 \\ \vdots \\ r u_n \end{pmatrix}.$$

§2 Punkte, Geraden und Ebenen

Vektorräume bieten einen guten Rahmen zur Behandlung geometrischer Fragen. Diese werden dabei *analytisch* behandelt, indem man die geometrischen Objekte rechnerisch 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, Pyramiden, ...).

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.

Ist X irgendein Punkt der Ebene oder des Raumes, etwa $X = (x_1 \mid x_2 \mid x_3)$, so betrachten wir den Verbindungsvektor vom Koordinatenursprung $O = (0 \mid 0 \mid 0)$ zu X :

$$\overrightarrow{OX} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}.$$

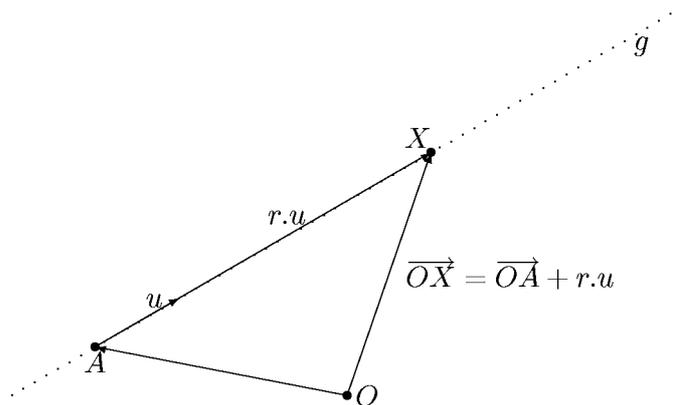
Diesen nennen wir den *Ortsvektor* \overrightarrow{OX} des Punktes X . Wie wir sehen, haben Punkt und Ortsvektor dieselben Koordinaten. Statt mit dem Punkt kann man also genausogut mit dem Ortsvektor arbeiten. Dies ermöglicht es, alle geometrischen Fragen vektoriell zu erfassen. (Man beachte aber, daß der Ortsvektor ein festgewähltes Koordinatensystem bzw. dessen Koordinatenursprung O erfordert.)

Auch beliebige Verbindungsvektoren zweier Punkte A, B kann man durch die Ortsvektoren \overrightarrow{OA} von A und \overrightarrow{OB} von B erfassen (siehe S. 4):

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

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 sei ein Punkt A und ein Richtungsvektor u einer Geraden. Wir bezeichnen mit $g(A, u)$ diese Gerade durch A mit u als Richtungsvektor. Wir wollen sämtliche Punkte X auf der Geraden algebraisch erfassen. Nun liegt X gerade dann auf der Geraden $g(A, u)$, 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. (Vergleiche Übung (1), Aufgabe 4)). Wir erhalten so für die Gerade $g(A, u)$ durch A mit Richtungsvektor u :

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

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

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

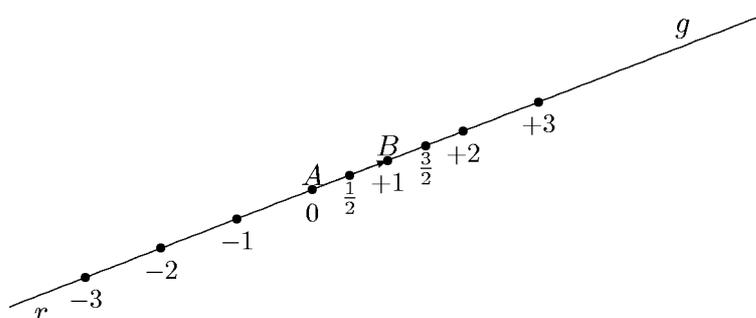
Man nennt diese beiden äquivalenten Beschreibungen eine *Parameterdarstellung* für die Gerade $g(A, u)$ durch den Punkt A mit dem Richtungsvektor u . Für jeden Wert des *Parameters* r liefert die Gleichung $\overrightarrow{OX} = \overrightarrow{OA} + r \cdot u$ den Ortsvektor eines Punktes X auf $g(A, u)$, 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 durch *einen* Parameter (=Zahl) erfaßt.

Ist nun g die Gerade durch zwei verschiedene Punkte A, B (wir bezeichnen sie entsprechend mit $g(A, B)$), so wählt man etwa A als den einen Punkt auf der Geraden g und $u = \overrightarrow{AB}$ als Richtungsvektor

($u \neq 0$ da $A \neq B$). Damit lautet eine Parameterdarstellung für die Gerade $g(A, B)$ durch zwei Punkte A, B :

$$X \in g(A, B) \iff \overrightarrow{OX} = \overrightarrow{OA} + r \cdot \overrightarrow{AB} \text{ für ein } r \in \mathbb{R}.$$

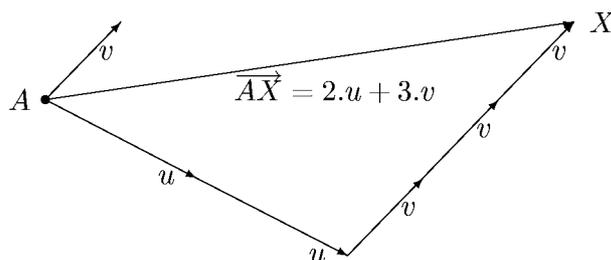
In der nachstehenden Skizze sind bei gegebenen Punkten A, B für einige Geradenpunkte die zugehörigen Parameterwerte angegeben. Man sieht: Der Geraden wird auf diese Weise eine Skala aufgeprägt und jeder Geradenpunkt wird durch einen Zahlwert (Parameter) erfaßt. Man benutzt dann zur Rechnung diese Parameterwerte 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 durch 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. Solche Vektoren nennt man *linear unabhängig*.

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

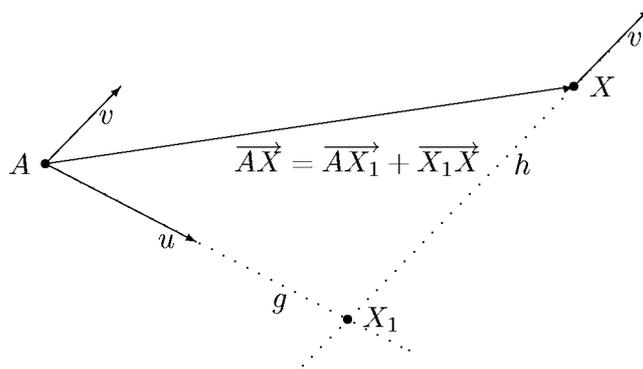


v , so verläßt man die Ebene nicht, man ‘kommt’ immer zu einem Endpunkt X , der in der Ebene liegt. Der dabei ‘zurückgelegte’ Vektor ist von der Form $r \cdot u + s \cdot v$. Dies bedeutet, daß alle Punkte X mit

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

notwendig zur Ebene gehören. (Man nennt solch einen Vektorterm $r \cdot u + s \cdot v$ mit $r, s \in \mathbb{R}$ eine *Vielfachsumme* oder auch eine *Linearkombination* von u und v .)

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 $r \cdot u$ von u und $\overrightarrow{X_1X}$ ein Vielfaches $s \cdot v$ von v . Also $\overrightarrow{AX} = \overrightarrow{AX_1} + \overrightarrow{X_1X} = r \cdot u + s \cdot v$.

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

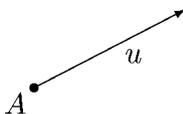
$$X \in e(A, u, v) \iff \overrightarrow{AX} = r \cdot u + s \cdot v \text{ für geeignete } r, s \in \mathbb{R}.$$

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

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

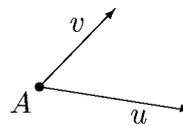
d. Affine Koordinatensysteme. Zur Beschreibung von Geraden und Ebenen haben wir Parameterdarstellungen benutzt. Diese wiederum wurden festgelegt durch Auswahl eines sog. *affinen Koordinatensystems*:

Bei Geraden:



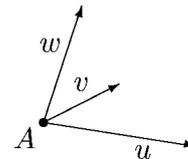
Ein Punkt A und ein Vektor $u \neq o$.

Bei Ebenen:



Ein Punkt A und zwei linear unabhängige Vektoren u, v .

Will man nun den gesamten (dreidimensionalen) *Raum* erfassen, so benötigt man wiederum einen Punkt A sowie nun *drei* Vektoren u, v, w , die wir wieder mit Anfangspunkt A darstellen. Um den ganzen Raum erfassen zu können, dürfen die Vektoren aber *nicht in einer Ebene liegen*. Dies bedeutet, daß *keiner eine Linearkombination der übrigen sein darf*.



§3 Lineare Gleichungssysteme und das Gauß'sche Eliminationsverfahren

a. Matrixschreibweise. Eine Gleichung heißt bekanntlich *linear*, wenn in ihr die Unbekannten nur in erster Potenz vorkommen. Ein *lineares* Gleichungssystem kann daher zunächst immer äquivalent in die folgende Form umgeformt werden:

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

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 und die rechten Seiten b_i sind dabei beliebige reelle Zahlen. Sie sind so bezeichnet, daß man an den Indizes (i bzw. j) erkennen kann, um welche Gleichung es sich handelt (erster Index i) und zu welcher Unbekannten der Koeffizient gehört (Index j).

Fixiert man die Reihenfolge der Unbekannten (x_1, x_2, x_3, \dots) , so kann man ein solches lineares Gleichungssystem vereinfacht in einer *Matrix* (einem rechteckigen Zahlenschema) notieren. Man schreibt dabei

die zu den einzelnen Unbekannten gehörenden Koeffizienten a_{ij} (samt Vorzeichen) in getrennten Spalten nebeneinander und nach einem Trennstrich ebenso die Zahlen b_i der rechten Seite. In dieser kompakten Notation erhält man dann für das obige lineare 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 Anzahl m der Gleichungen finden wir nun als Zahl der *Zeilen* und die Anzahl n der Unbekannten als Zahl der *Spalten der Koeffizientenmatrix* wieder.

b. Gauß-Elimination und der Rang. Lineare Gleichungssysteme löst man mit Hilfe des *Gauß'schen Eliminationsverfahrens*. Die dabei benutzten Umformungen sind

1. Zeilentausch,
2. Multiplikation einer Zeile mit einer Zahl $\neq 0$,
3. Addition eines beliebigen reellen Vielfachen einer Zeile zu einer anderen,
4. Spaltentausch (mit Beachtung der Zuordnung von Spalten und Unbekannten).

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

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

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

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 n) sind (falls überhaupt vorhanden) Nullzeilen.

Die hierbei auftretende Zahl r nennen wir den *Rang* des Gleichungssystems oder der Koeffizientenmatrix. Es gilt $r \leq n$ und $r \leq m$.

Der Rang r einer Matrix ist die Anzahl der 'Nicht-Null-Zeilen' in der Matrix nach *vollständig* durchgeführter Gauß-Elimination.

Wir vermerken ohne Beweis: Der Rang ist unabhängig von den gewählten Umformungsschritten.

c. Lösbarkeit und Lösungsvielfalt. Die $m - r$ Nullzeilen am Ende stellen Gleichungen der Form

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

dar, wobei b'_{r+1}, \dots, b'_m die bei der Gauß-Elimination veränderten Werte der rechten Seite sind. Diese entscheiden über die Lösbarkeit des gesamten Gleichungssystems:

1. Ist *eine* dieser Gleichungen *nicht erfüllt*, d. h. ist einer der Werte b'_{r+1}, \dots, b'_m *ungleich* 0, so ist das gesamte Gleichungssystem *unlösbar*.
2. Sind *alle* diese Gleichungen *erfüllt*, d. h. sind alle $m - r$ Werte b'_{r+1}, \dots, b'_m *gleich* 0, 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*.

Ein lineares Gleichungssystem mit m Gleichungen und dem Rang r ist genau dann lösbar, wenn alle $m - r$ Lösbarkeitsbedingungen erfüllt sind: $b'_{r+1} = b'_{r+2} = \dots = b'_m = 0$.

Im Falle der Lösbarkeit kann man die letzten $n - r$ Unbekannten x_{r+1}, \dots, x_n frei wählen und dann die ersten r Gleichungen nach x_r, \dots, x_1 auflösen. Auf diese Weise kann man beliebig gewählte x_{r+1}, \dots, x_n immer zu einer Lösung x_1, \dots, x_n vervollständigen. Man erhält so eine Parameterdarstellung für die Lösungsmenge, wobei die $n - r$ Unbekannten x_{r+1}, \dots, x_n die frei wählbaren Parameter sind und die zugehörigen Vektoren (siehe Übungen) sich als linear unabhängig erweisen. Damit wird die Lösungsmenge eine Gerade (bei $n - r = 1$), eine Ebene (bei $n - r = 2$) bzw. allgemein eine $n - r$ -dimensionale affine Teilmenge. Da im Falle der Unlösbarkeit die Lösungsmenge leer ist, kann man sagen:

Die Lösungsmenge eines linearen Gleichungssystems mit n Unbekannten und dem Rang r ist leer oder eine $n - r$ -dimensionale affine Menge.

d. Spezialfälle. 1. Einen Sonderfall stellen die sog. *homogenen* Gleichungssysteme dar. Dies sind solche, bei denen die rechte Seite nur Nullen enthält. Für diese homogenen Systeme gelten die obigen Überlegungen selbstverständlich auch, nur gilt *zusätzlich*:

Ein homogenes lineares Gleichungssystem 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:

Die Lösungsmenge eines homogenen Gleichungssystems mit n Unbekannten und dem Rang r ist eine $n - r$ -dimensionale affine Teilmenge, die durch den Koordinatenursprung verläuft.

2. Im Fall $r = m$ gibt es keine Nullzeilen und daher auch *keine* Bedingungen für die Lösbarkeit; das lineare Gleichungssystem ist lösbar! Und dies gilt unabhängig von den Werten auf der rechten Seite des Gleichungssystems.

Ist der Rang r eines linearen Gleichungssystems *gleich* der Anzahl m der Gleichungen, so ist das Gleichungssystem *universell* (d. h. für jede beliebige rechte Seite) lösbar.

3. Im Fall $n = r$ gibt es keine frei wählbaren Parameter, die Lösung ist also eindeutig (wenn überhaupt eine existiert).

Ist der Rang r eines linearen Gleichungssystems *gleich* der Anzahl n der Unbekannten, so hat das Gleichungssystem keine oder genau eine Lösung.

4. Die Kombination der beiden vorangehenden Spezialfälle führt zu dem folgenden 'Idealfall' linearer Gleichungssysteme:

Hat ein lineares Gleichungssystem genau so viele Gleichungen wie Unbekannte und den größtmöglichen Rang $r = n = m$, so ist es bei *jeder* beliebigen rechten Seite *eindeutig* lösbar.

§4 Gleichungsdarstellung affiner Teilmengen

a. Parameterdarstellungen. Bisher haben wir affine Teilmengen M (Geraden, Ebenen, ...) immer durch Parameterdarstellungen

$$X \in M \iff x = a + r_1 u_1 + \dots + r_d u_d \quad (r_i \in \mathbb{R})$$

beschrieben. Dabei bezeichnet x den Ortsvektor des beliebigen Punktes X , a den Ortsvektor eines fest gewählten Punktes A in der affinen Teilmenge M 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 $d = 2$, ...).

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 Punkte 'aufzählen', die zu dem geometrischen Gebilde gehören. Es ist jedoch schwierig, von einem beliebig vorgelegten Punkt zu *entscheiden*, ob er zu der affinen Teilmenge gehört oder nicht.

b. Ein Beispiel. Gegeben eine Ebene e mit der Parameterdarstellung

$$x = a + ru + sv = \begin{pmatrix} 2 \\ -1 \\ 0 \end{pmatrix} + r \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix} + s \begin{pmatrix} 0 \\ 1 \\ 2 \end{pmatrix} \quad (r, s \in \mathbb{R})$$

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

Um dies etwa für P zu entscheiden, muß man untersuchen, ob der Ortsvektor p von P eine Darstellung in Form der Parameterdarstellung hat, d. h. ob es Parameterwerte r, s 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.

Statt mehrfach dieselben Rechnungen durchzuführen, kann man das Problem allgemein lösen. Man untersucht für einen beliebigen Punkt $X = (x \mid y \mid z)$, ob er zu e gehört oder nicht. Man muß also das lineare Gleichungssystem mit folgender erweiterter Matrix

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

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:

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

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:

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

Dies ist nun eine für jeden Punkt X einfach zu überprüfende Bedingung. Zum Beispiel für den Punkt $P = (1 \mid 4 \mid 5)$ ist diese Gleichung erfüllt: $-5 \cdot 1 - 2 \cdot 4 + 5 + 8 = 0$; P liegt also in der Ebene. Für $Q = (3 \mid -2 \mid 1)$ 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 Gleichung.

c. Gleichungsdarstellung affiner Teilmengen. Bei der Lösung linearer Gleichungssysteme bestimmt man für die Lösungsmenge eine Parameterdarstellung. Im obigen Beispiel b. ging man umgekehrt vor: 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 $X = (x \mid y \mid z)$ zu der Geraden g durch die Punkte $A = (2 \mid 1 \mid -1)$ und $B = (-1 \mid 1 \mid 0)$?

Man bestimmt zunächst eine Parameterdarstellung für g :

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \\ -1 \end{pmatrix} + r \begin{pmatrix} -3 \\ 0 \\ 1 \end{pmatrix} \quad (r \in \mathbb{R})$$

und untersucht diese dann als Gleichungssystem mit nur einer Unbekannten (r) auf Lösbarkeit. Die erweiterte Matrix ist

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

und 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 $m - r = 3 - 1 = 2$: Beide Gleichungen $y - 1 = 0$ und $3(z + 1) + (x - 2) = x + 3z + 1 = 0$ müssen erfüllt sein, damit der Punkt $X = (x \mid y \mid z)$ auf der Geraden g liegt. g ist damit Lösungsmenge des linearen Gleichungssystems

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

Die Methode, von einer Parameterdarstellung einer affinen Menge zu einer Gleichungsdarstellung zu kommen, kann man so zusammenfassen:

Man faßt die Parameterdarstellung als ein lineares Gleichungssystem mit den Parametern als Unbekannten auf und untersucht dieses mittels Gauß-Elimination auf Lösbarkeit. Die Lösbarkeitsbedingungen, die sich auf den rechten Seiten der Nullzeilen ergeben, stellen die gesuchten linearen Gleichungen dar.

Bereits die beiden obigen Beispiele zeigen einen 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 in einem m -dimensionalen Raum ist Lösungsmenge eines linearen Gleichungssystem bestehend aus $m - d$ Gleichungen.

Eine Begründung dieser Tatsache ergibt sich im nachfolgenden Abschnitt §5.

§5 Lineare Unabhängigkeit

a. Parameterdarstellungen und lineare Unabhängigkeit. Zur Beschreibung von affinen Mengen (Gerade, Ebene, Raum) benutzt man Parameterdarstellungen, die sich aus affinen Koordinatensystemen (siehe §2 d.) für diese Mengen ergeben. Ein solches affines Koordinatensystem besteht aus einem Punkt A und einem oder mehreren Vektoren.

Bei einer Geraden benötigt man *einen* Vektor u und dieser muß vom Nullvektor verschieden sein; bei einer Ebene benötigt man *zwei* Vektoren u, v und diese dürfen keine Vielfachen voneinander sein; um den ganzen Raum zu erfassen, benötigt man *drei* Vektoren u, v, w , von denen keiner eine Linearkombination der übrigen sein darf.

In allen drei o. g. Fällen nennt man die auftretenden Vektoren jeweils *linear unabhängig*. Wir wollen nun eine *einheitliche* Definition dieses Begriffes geben. Diese ist zugleich für beliebige Vektorräume gültig.

Definition: Es sei V ein Vektorraum und u_1, \dots, u_n irgendwelche Vektoren in V .

a) Eine *Linearkombination* von u_1, \dots, u_n ist eine Vielfachsumme dieser Vektoren, d. h. ein Term der Form:

$$r_1 u_1 + r_2 u_2 + \dots + r_n u_n \quad \text{mit Koeffizienten } r_i \in \mathbb{R}.$$

b) Wir nennen ganz allgemein ein System von n Vektoren u_1, \dots, u_n eines Vektorraums V *linear unabhängig*, wenn keiner der Vektoren u_i sich als Linearkombination der *übrigen* darstellen läßt.

b. Koeffizientenvergleich. Die entscheidenden Sachverhalte über linear unabhängige Vektoren liegen in den folgenden Charakterisierungen:

Satz. Für n Vektoren u_1, \dots, u_n in einem Vektorraum V sind die folgenden Aussagen äquivalent:

i) (Lineare Unabhängigkeit)

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

ii) (Nur die triviale Nulldarstellung)

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

$$r_1 u_1 + r_2 u_2 + \dots + r_n u_n = o \iff r_1 = r_2 = \dots = r_n = 0.$$

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

$$r_1 u_1 + \dots + r_n u_n = s_1 u_1 + \dots + s_n u_n \iff r_1 = s_1, \dots, r_n = s_n.$$

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 der *linear unabhängigen* Vektoren u_1, \dots, u_n denselben Vektor dar, so ist dies nur möglich, wenn jeweils die einander entsprechenden Koeffizienten übereinstimmen. Auf diese Weise ist eine Vektorgleichung äquivalent zu n Gleichungen für Zahlen.

Die Eigenschaft ii) ist folgendermaßen zu verstehen: Aus n beliebigen Vektoren u_1, \dots, u_n kann man immer eine Linearkombination bilden, die den Nullvektor o darstellt; man wählt alle Koeffizienten $r_i = 0$:

$$0 \cdot u_1 + \dots + 0 \cdot u_n = o.$$

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 $r_1 u_1 + \dots + r_n u_n = o$ stellt eine echte *Relation* (Beziehung) zwischen den Vektoren dar, wenn wenigstens ein Koeffizient $r_i \neq 0$ ist. Man kann also sagen, daß n 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*: $o = r_1 u_1 + \dots + r_n u_n$. 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 u_1 auflösen:

$$o = r_1 u_1 + \dots + r_n u_n \iff r_1 u_1 = -r_2 u_2 - \dots - r_n u_n, \iff u_1 = -\frac{r_2}{r_1} u_2 - \dots - \frac{r_n}{r_1} u_n.$$

Damit wäre aber u_1 eine Linearkombination der übrigen Vektoren u_2, \dots, u_n , was nach i) nicht sein kann. Genauso folgert man einen Widerspruch, wenn ein anderer Koeffizient $r_i \neq 0$ ist.

ii) \implies iii): Unter Verwendung von ii) erhalten wir folgende Äquivalenzen:

$$\begin{aligned} r_1 u_1 + \dots + r_n u_n = s_1 u_1 + \dots + s_n u_n &\iff (r_1 - s_1)u_1 + \dots + (r_n - s_n)u_n = 0 \\ &\iff_{ii)} r_1 - s_1 = \dots = r_n - s_n = 0 \\ &\iff r_1 = s_1, \dots, r_n = s_n \end{aligned}$$

Damit ist iii) bewiesen.

iii) \implies i): Wäre i) falsch, etwa u_1 eine Linearkombination der übrigen:

$$u_1 = r_2 u_2 + \dots + r_n u_n,$$

so erhielte man

$$1 \cdot u_1 + 0 \cdot u_2 + \dots + 0 \cdot u_n = 0 \cdot u_1 + r_2 u_2 + \dots + r_n u_n.$$

Nach iii) folgte dann aber $1=0$ (Koeffizient von u_1 auf der linken Seite = Koeffizient von u_1 auf der rechten Seite), offenbar ein Widerspruch. Also kann u_1 keine Linearkombination der anderen Vektoren sein. Genauso schließt man, wenn ein anderes u_i eine Linearkombination der übrigen Vektoren wäre.

c. Geometrische Anwendungen. Eine besondere Bedeutung hat der Koeffizientenvergleich linear unabhängiger Vektoren beim Nachweis *allgemeiner* geometrischer Zusammenhänge.

Beispiel: In jedem beliebigen Dreieck ABC gilt:

1. Je zwei Seitenhalbierende schneiden sich in einem Punkt; dieser teilt die Seitenhalbierenden im Verhältnis $2 : 1$, wobei der größere Teil beim Eckpunkt liegt.
2. Alle drei Seitenhalbierenden schneiden sich in demselben Punkt, der *Schwerpunkt* des Dreiecks genannt wird.
3. Für den Schwerpunkt S gilt die folgende Berechnungsformel:

$$\vec{OS} = \frac{1}{3}(\vec{OA} + \vec{OB} + \vec{OC})$$

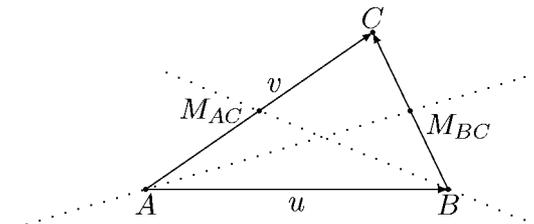
für jeden beliebigen Punkt O .

Beweis: Da die Eckpunkte des Dreiecks nicht konkret gegeben sind, kann es nicht darum gehen, die Koordinaten des Schnittpunktes zu ermitteln. Vielmehr geht es um die Lage des Schwerpunktes *relativ* zu den Eckpunkten. Nun bilden drei Punkte A, B, C genau dann ein *Dreieck*, wenn sie nicht auf einer Geraden liegen (nicht *kollinear* sind). Dies ist genau dann der Fall, wenn die Vektoren $u = \vec{AB}$ und $v = \vec{AC}$ keine Vielfachen voneinander sind, wenn u, v also linear unabhängig sind. Damit bestimmt ein Dreieck ein affines Koordinatensystem für die Ebene $e = e(A, B, C)$ mit A als Koordinatenursprung und u, v als Richtungsvektoren. Dieses wird bei allen weiteren Rechnungen zugrundegelegt.

ad 1. Wir untersuchen die Seitenhalbierenden $g(A, M_{BC})$ und $g(B, M_{AC})$. Dazu stellen wir zunächst Parameterdarstellungen dafür auf (mit A als Koordinatenursprung):

$$\begin{aligned} X \in g(A, M_{BC}) &\iff \vec{AX} = r \cdot \vec{AM}_{BC} & (r \in \mathbb{R}). \\ X \in g(B, M_{AC}) &\iff \vec{AX} = \vec{AB} + s \cdot \vec{BM}_{AC} & (s \in \mathbb{R}). \end{aligned}$$

Als nächsten Schritt stelle man in diesen Parameterdarstellungen die rechten Seiten als *Linearkombinationen* der zugrundegelegten linear unabhängigen Vektoren u, v dar. Dabei orientiere man sich an einer kleinen Skizze wie rechts.



$$\begin{aligned} g(A, M_{BC}) : \quad \vec{AX} &= r \vec{AM}_{BC} = r \left(\vec{AB} + \frac{1}{2} \vec{BC} \right) \\ &= r \left(u + \frac{1}{2} (\vec{BA} + \vec{AC}) \right) = r \left(u + \frac{1}{2} (-u + v) \right) \\ &= \frac{r}{2} \cdot u + \frac{r}{2} \cdot v. \\ g(B, M_{AC}) : \quad \vec{AX} &= \vec{AB} + s \cdot \vec{BM}_{AC} = u + s \left(\vec{BA} + \frac{1}{2} \vec{AC} \right) \\ &= u + s \left(-u + \frac{1}{2} v \right) = (1 - s) \cdot u + \frac{s}{2} \cdot v. \end{aligned}$$

Nun kann man den Schnittpunkt bestimmen, indem man wie üblich die Parameterdarstellungen der beiden Geraden ‘gleichsetzt’:

$$g(A, M_{BC}) \cap g(B, M_{AC}) : \quad \frac{r}{2} \cdot u + \frac{r}{2} \cdot v = (1-s) \cdot u + \frac{s}{2} \cdot v.$$

Da u, v linear unabhängig sind, kann man in dieser Vektorgleichung nun *Koeffizientenvergleich* durchführen. Dieser besagt, daß die Koeffizienten von u auf beiden Seiten der Gleichung übereinstimmen müssen: $r/2 = 1-s$, und genauso die Koeffizienten von v : $r/2 = s/2$. Durch den Koeffizientenvergleich wird also die Vektorgleichung überführt in ein lineares Gleichungssystem für r, s ! Dieses löst man nun wie gewohnt:

$$\begin{bmatrix} \frac{r}{2} = 1-s \\ \frac{r}{2} = \frac{s}{2} \end{bmatrix} \iff \begin{bmatrix} r = 1-s \\ \frac{r}{2} = 1-r \end{bmatrix} \iff \begin{bmatrix} r = s \\ r = \frac{2}{3} \end{bmatrix} \iff r = s = \frac{2}{3}.$$

Damit ist gezeigt, daß die Seitenhalbierenden sich in einem Punkt S schneiden, und daß dieser die Seitenhalbierenden im Verhältnis $\frac{2}{3} : \frac{1}{3} = 2 : 1$ teilt. Der größere Abschnitt liegt dabei auf Seiten der jeweiligen Ecke.

ad 2. Da Aussage 1. allgemeingültig (nicht auf spezielle Punkte beschränkt) ist, muß auch die dritte Seitenhalbierende die anderen im genannten Verhältnis teilen, also in demselben Punkt S schneiden: Alle *drei* Seitenhalbierenden schneiden sich in einem Punkt.

ad 3. Den Schnittpunkt S erhält man, indem man einen der gefundenen Parameterwerte in die zugehörige Parameterdarstellung einsetzt:

$$\overrightarrow{AS} = \frac{2/3}{2} \cdot u + \frac{2/3}{2} \cdot v = \frac{1}{3}(\overrightarrow{AB} + \overrightarrow{AC}).$$

Indem man auf beiden Seiten den Vektor \overrightarrow{PA} addiert, erhält man:

$$\overrightarrow{PA} + \overrightarrow{AS} = \overrightarrow{PA} + \frac{1}{3}(\overrightarrow{AB} + \overrightarrow{AC}) = \frac{1}{3}(3 \cdot \overrightarrow{PA} + \overrightarrow{AB} + \overrightarrow{AC}) = \frac{1}{3}(\overrightarrow{PA} + \overrightarrow{PA} + \overrightarrow{AB} + \overrightarrow{PA} + \overrightarrow{AC}) = \frac{1}{3}(\overrightarrow{PA} + \overrightarrow{PB} + \overrightarrow{PC})$$

und damit die behauptete Formel für \overrightarrow{PS} . Weitere Anwendungen siehe Übungen (6).

d. Rang und lineare Unabhängigkeit. Sind u_1, \dots, u_n Spaltenvektoren, 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

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

die Vektoren u_i als Spalten hat. Nun hat ein homogenes lineares Gleichungssystem immer eine Lösung, und diese ist eindeutig, wenn der Rang r gleich der Zahl n der Unbekannten ist (siehe die Resultate über lineare Gleichungssysteme). Dies bedeutet:

n Vektoren sind genau dann linear unabhängig, wenn die mit ihnen als Spaltenvektoren gebildete Matrix den Rang n hat.

Auf dieser Tatsache basiert die Begründung für die Beziehung zwischen Dimension einer affinen Menge und Anzahl der zu ihrer Beschreibung benötigten Gleichungen (siehe Ende von §4). Wir wollen diese hier nachtragen.

Gegeben ist eine Parameterdarstellung $x = a + r_1 u_1 + \dots + r_d u_d$ einer d -dimensionalen affinen Teilmenge im m -dimensionalen Raum ($m = 2, 3, \dots$). Ein Punkt $X = (x_1, \dots, x_m)$ mit Ortsvektor x gehört zu dieser affinen Menge, wenn das Gleichungssystem mit der erweiterten Matrix

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

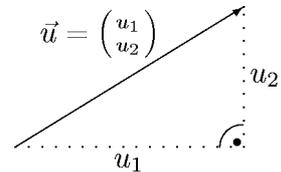
lösbar ist. Dieses System hat die Zeilenzahl m und die Spaltenzahl d . Da die Spaltenvektoren u_1, \dots, u_d linear unabhängig sind, hat es den maximal möglichen Rang d (siehe oben). Damit erhält man nach der Gauß-Elimination $m - d$ Bedingungen für die Lösbarkeit. Diese Bedingungen sind Gleichungen, 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.

§6 Länge, Skalarprodukt, Orthogonalität

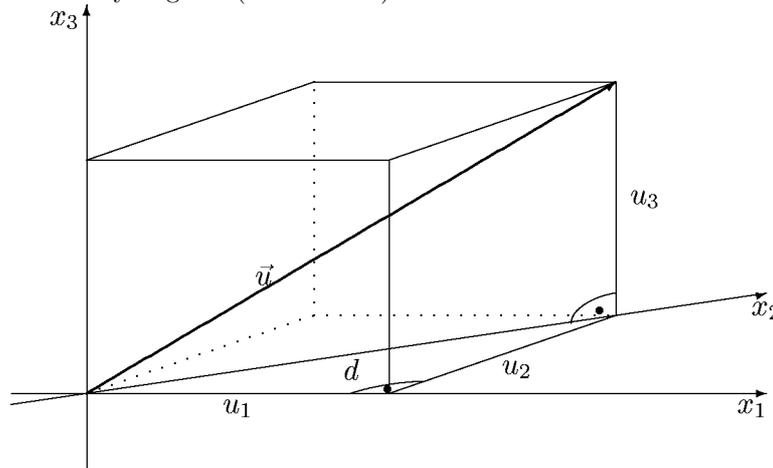
Wir haben bisher analytische Geometrie allein auf dem Vektorbegriff aufgebaut. Wir haben affine Mengen, ihre gegenseitige Lage, ihre Schnittmengen und Längenverhältnisse betrachtet, nicht jedoch Längen absolut. Ebenso haben wir den Winkelbegriff in unserem bisherigen Geometrieaufbau nicht behandelt. 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 $\vec{u} = \begin{pmatrix} u_1 \\ u_2 \end{pmatrix}$ in der Ebene gegeben, so berechnet man seine Länge $|\vec{u}|$ leicht mit Hilfe des Satzes des Pythagoras:

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



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 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 \vec{u} die Länge von \vec{u}

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

Aufgrund dieser Formel liegt die folgende allgemeine Definition nahe:

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

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

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

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

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 $\vec{u} \cdot \vec{v}$ durch

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

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

$$|\vec{u}|^2 = \vec{u} \cdot \vec{u} :$$

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 $\vec{u}, \vec{v}, \vec{w}$ und beliebige reelle Zahlen r :

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

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!

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

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:

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

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

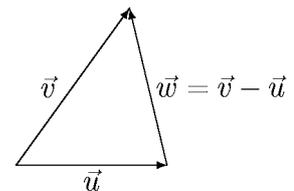
$$(\vec{u} + \vec{v}) \cdot (\vec{u} - \vec{v}) = \vec{u} \cdot \vec{u} - \vec{v} \cdot \vec{v} = |\vec{u}|^2 - |\vec{v}|^2.$$

Diese Formeln haben vielfältige geometrische Konsequenzen (siehe in den Übungen).

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

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

Man erkennt eine Ähnlichkeit mit dem Satz des Pythagoras, nur daß hier bei beliebigen Dreiecken ein *Korrekturterm* auftritt, nämlich $-2\vec{u} \cdot \vec{v}$. Man sieht daran:



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

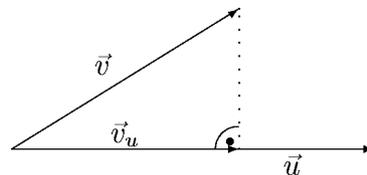
Definition: Zwei Vektoren des \mathbb{R}^n nennt man *orthogonal*, in Zeichen $\vec{u} \perp \vec{v}$, wenn ihr Skalarprodukt Null ist:

$$\vec{u} \perp \vec{v} \iff \vec{u} \cdot \vec{v} = 0.$$

Man beachte, daß bei dieser Definition der Nullvektor \vec{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 \vec{v} auf einen anderen $\vec{u} \neq \vec{o}$ aufzeigen. Dabei ist der Projektionsvektor \vec{v}_u (siehe Skizze) charakterisiert durch:

$$\vec{v}_u = r\vec{u} \text{ ist Vielfaches von } u \quad \text{und} \quad \vec{v} - \vec{v}_u \perp \vec{u}.$$



Mit den Eigenschaften des Skalarproduktes kann man daraus sehr leicht $\vec{v}_u = r\vec{u}$ berechnen, denn:

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

Diese Gleichung läßt sich (wegen $|\vec{u}|^2 \neq 0$) nach r auflösen und ergibt die folgende allgemeine Formel für r und damit \vec{v}_u :

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

In mehr geometrischer Sprechweise ist dies die Bestimmung eines *Lotfußpunktes*: Man betrachtet die Richtung von \vec{u} als den 'Boden' und läßt von der Spitze von \vec{v} ein *Lot* herunterhängen: Die Spitze von \vec{v}_u ist dann der Lotfußpunkt. (Zur Berechnung von Lotfußpunkten und ihre Bedeutung siehe Übungen (8).)

Der Begriff der Orthogonalität ist nicht nur auf Vektoren, sondern auch auf *affinen Mengen* anwendbar. So nennt man zwei *Geraden* g, g' orthogonal zueinander, wenn sie Richtungsvektoren $\vec{u} \neq \vec{o}$ bzw. $\vec{u}' \neq \vec{o}$ besitzen, die orthogonal zueinander sind. Dies bedeutet dann aber, daß *jeder* Richtungsvektor der einen Gerade orthogonal ist zu *jedem* Richtungsvektor der anderen:

$$g \perp g' \iff \overrightarrow{AB} \perp \overrightarrow{A'B'} \quad \text{für alle } A, B \in g, A', B' \in g'. \quad (*)$$

(Begründung: $\overrightarrow{AB} = r\vec{u}$, $\overrightarrow{A'B'} = r'\vec{u}'$, also $\overrightarrow{AB} \cdot \overrightarrow{A'B'} = (r\vec{u}) \cdot (r'\vec{u}') = rr' \cdot \vec{u} \cdot \vec{u}' = 0$, da $\vec{u} \perp \vec{u}'$, also $\vec{u} \cdot \vec{u}' = 0$.)

Für eine Gerade g und eine Ebene e bedeutet Orthogonalität, daß ein Richtungsvektor $\vec{u} \neq \vec{o}$ von g orthogonal zu *zwei linear unabhängigen* Richtungsvektoren \vec{v}, \vec{w} der Ebene e sein muß. Dies hat dann wieder die Eigenschaft (*) (für g, e) zur Folge.

Die Orthogonalität von Ebenen definiert man über die Orthogonalität ihrer Normalenvektoren (siehe d.).

d. Normalenvektoren und Gleichungen. Ein Vektor, der orthogonal ist zu einer affinen Menge (Gerade, Ebene, ...), nennt man einen *Normalenvektor*. Diese haben eine enge Beziehung zu den Gleichungen, die die affine Menge beschreiben.

Es sei e eine Ebene im dreidimensionalen Raum und $ax + by + cz + d = 0$ eine Gleichung für e , d. h. e besteht aus allen Punkten, deren Koordinaten diese Gleichung erfüllen:

$$X = (x \mid y \mid z) \in e \iff ax + by + cz + d = 0.$$

Faßt man nun die Koeffizienten a, b, c der Unbekannten in der gegebenen Gleichung zu einem Vektor zusammen:

$$\vec{n} = \begin{pmatrix} a \\ b \\ c \end{pmatrix},$$

so kann man damit die Ebenengleichung in folgender Form schreiben:

$$X \in e \iff \vec{n} \cdot \vec{x} + d = 0. \quad (*)$$

Hieraus ergibt sich, daß $\vec{n} \neq \vec{o}$ ein Normalenvektor zu e ist, also orthogonal ist zu *jedem* Verbindungsvektor \overrightarrow{PQ} zweier Punkte $P, Q \in e$:

$$\vec{n} \perp \overrightarrow{PQ} \quad \text{für alle } P, Q \in e.$$

Beweis: $\vec{n} \neq \vec{o}$, denn sonst wäre $a = b = c = 0$ und $ax + by + cz + d = 0$ wäre keine Ebenengleichung. Seien nun $P = (p_1 \mid p_2 \mid p_3)$ und $Q = (q_1 \mid q_2 \mid q_3)$ zwei beliebige Punkte der Ebene e . Dann erfüllen sie beide die Gleichung (*) für e , also gilt:

$$\vec{n} \cdot \vec{p} + d = 0 \quad \text{und} \quad \vec{n} \cdot \vec{q} + d = 0.$$

Durch Subtraktion beider Gleichungen erhält man

$$0 = \vec{n} \cdot \vec{q} - \vec{n} \cdot \vec{p} = \vec{n} \cdot (\vec{q} - \vec{p}) = \vec{n} \cdot \overrightarrow{PQ}, \quad \text{also} \quad \vec{n} \perp \overrightarrow{PQ}.$$

Da \vec{n} ein Normalenvektor der Ebene e ist, nennt man (*) eine *Normalengleichung* von e .

Was hier für Ebenen e im 3-dimensionalen Raum gesagt wurde, gilt sinngemäß genauso für Geraden g in der Ebene: Ist $ax + by + d = 0$ eine Gleichung für eine Gerade g in der Ebene, so ist der Vektor $\vec{n} = \begin{pmatrix} a \\ b \end{pmatrix}$ ein Normalenvektor zu g : \vec{n} ist senkrecht zu einem (und damit jedem) Richtungsvektor von g . Diese Tatsache kann man sich zunutze machen, um auf einfache Weise Gleichungen für *Geraden in Ebenen* zu bestimmen (siehe späteres Beispiel).

e. Abstandsberechnungen. Ihre wichtigste Bedeutung haben Normalenvektoren bei der Berechnung von Abständen. 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.)

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 auf eine affine Menge bestimmen.

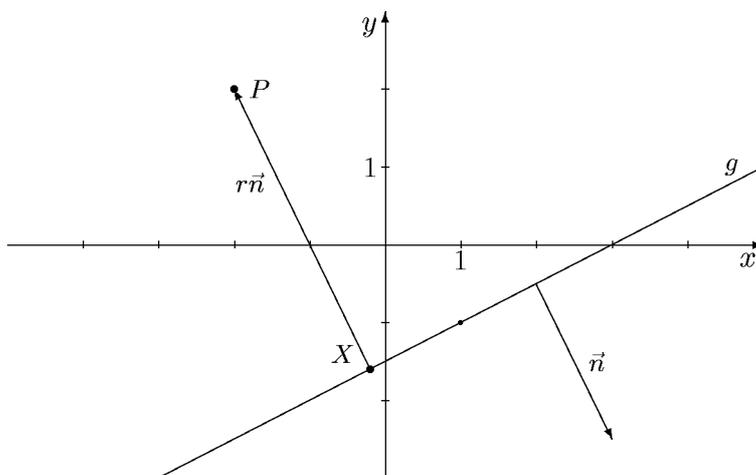
Wir wollen nun aber eine Methode kennenlernen, mit der man den Abstand bestimmen kann, *ohne* den Lotfußpunkt zu ermitteln. Diese ist jedoch nicht allgemein anwendbar, sondern nur auf Geraden in der Ebene und auf Ebenen im Raum. Entscheidend ist, daß in beiden Fällen die affine Menge durch *eine* lineare Gleichung beschrieben werden kann.

Hesse'sche Abstandsformel: Wir gehen aus von *einer* linearen Gleichung: $ax + by + d = 0$ (Lösungsmenge eine Gerade g im \mathbb{R}^2) bzw. $ax + by + cz + d = 0$ (Lösungsmenge eine Ebene e im dreidimensionalen Raum \mathbb{R}^3). Dann berechnet man den Abstand eines beliebigen Punktes P von g bzw. e wie folgt:

$$d(P, g) = \frac{|ap_1 + bp_2 + d|}{\sqrt{a^2 + b^2}} \quad \text{bzw.} \quad d(P, e) = \frac{|ap_1 + bp_2 + cp_3 + d|}{\sqrt{a^2 + b^2 + c^2}}.$$

Man muß also die Koordinaten des gegebenen Punktes P in die Geraden-/Ebenengleichung einsetzen und dann den Betrag der entstehenden Zahl durch den angegebenen Wurzelterm dividieren. Dieser Wurzelterm ist gerade die Länge des aus der Gleichung unmittelbar ablesbaren Normalenvektors.

Wir wollen dieses Resultat jetzt für den zweiten Fall begründen. Es sei also e eine Ebene (von der in der nachstehenden Skizze nur die Schnittgerade g mit der Zeichenebene sichtbar ist). Ist $ax + by + cz + d = 0$



eine Gleichung für e , so ist $\vec{n} = \begin{pmatrix} a \\ b \\ c \end{pmatrix}$ ein Normalenvektor und die Gleichung nimmt folgende Gestalt an:

$$X \in e \iff \vec{n} \cdot \vec{x} + d = 0 \quad (*)$$

Um den Abstand eines Punktes P von der Ebene e zu bestimmen, sucht man zunächst den Fußpunkt $X \in e$ des Lotes von P auf e ; der gesuchte Abstand ist dann die Länge des Vektors \overline{XP} . Wegen $X \in e$ gilt $\vec{n} \cdot \vec{x} + d = 0$ und wegen $\overline{XP} \perp e$ muß gelten $\overline{XP} = r \cdot \vec{n}$ (siehe die zweidimensionale Skizze mit einer Geraden g statt einer Ebene e). Wir berechnen nun zunächst r :

$$\begin{aligned} \overline{XP} = r \cdot \vec{n} &\implies \overline{XP} \cdot \vec{n} = r \cdot \vec{n} \cdot \vec{n} \iff (\vec{p} - \vec{x}) \cdot \vec{n} = r \cdot |\vec{n}|^2 \\ &\iff \vec{p} \cdot \vec{n} - \vec{x} \cdot \vec{n} = r \cdot |\vec{n}|^2 \stackrel{(*)}{\iff} \vec{p} \cdot \vec{n} + d = r \cdot |\vec{n}|^2 \end{aligned}$$

Mit dem sich ergebenden Wert für r

$$r = \frac{\vec{n} \cdot \vec{p} + d}{|\vec{n}|^2} = \frac{ap_1 + bp_2 + cp_3 + d}{a^2 + b^2 + c^2}$$

bestimmen wir den Abstand

$$d(P, e) = |\overline{XP}| = |r| \cdot |\vec{n}| = \frac{|\vec{n} \cdot \vec{p} + d|}{|\vec{n}|} = \frac{|ap_1 + bp_2 + cp_3 + d|}{\sqrt{a^2 + b^2 + c^2}}.$$

Dies ist die oben behauptete Formel.

Man beachte, daß mit der expliziten Formel für r (die übrigens keinen Wurzelterm enthält) auch der Lotfußpunkt schnell berechenbar wird. Außerdem zeigt das Vorzeichen von r an, auf welcher Seite der Ebene der Punkt P liegt: Ist $r > 0$, so liegt P auf der Seite von e , in die der Normalenvektor \vec{n} weist, andernfalls auf der entgegengesetzten Seite.

Was hier für Ebenen e im Raum \mathbb{R}^3 durchgeführt wurde, gilt sinngemäß genauso für Geraden g in der Ebene \mathbb{R}^2 . Wir wollen dies einmal beispielhaft an der skizzierten ebenen Situation durchführen. Wir zeigen dabei zugleich den Nutzen der Normalenvektoren bei der Bestimmung einer Gleichung für Geraden in der Ebene.

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

2. Ein Normalenvektor ist daher $\vec{n} = \begin{pmatrix} 1 \\ -2 \end{pmatrix}$. (Allgemein ist der Vektor $\begin{pmatrix} b \\ -a \end{pmatrix}$ orthogonal zu $\begin{pmatrix} a \\ b \end{pmatrix}$, denn das Skalarprodukt ergibt $ba - ab = 0$.)

3. Man setzt daher eine Gleichung für g in der Form $1 \cdot x - 2 \cdot y + d = 0$ an. d bestimmt man, indem man einen Geradenpunkt (etwa $(3|0)$) einsetzt: $3 - 2 \cdot 0 + d = 0$. Dies ergibt $d = -3$ und $x - 2y - 3 = 0$ 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:

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