

Mathematik für das Lehramt an der Grund- und
Mittelstufe sowie an Sonderschulen
Teil II (Lineare Algebra und analytische Geometrie): SoSe
07

Bodo Werner
<mailto:werner@math.uni-hamburg.de>

12. Juli 2007

Inhaltsverzeichnis

II	Lineare Algebra und analytische Geometrie	7
1	Der Anschauungsraum und seine n-dimensionale Form	9
1.1	Einführung	9
1.1.1	Rückblick auf Mathe I	12
1.2	Vektoren	15
1.2.1	Punkte in der Anschauungsebene und im Anschauungsraum	15
1.2.2	Vektoren als „Pfeile“	17
1.2.3	Vektoren im n-dimensionalen Raum	19
1.2.4	Vektoren in der Statistik	20
1.3	Vektoraddition	21
1.3.1	Geometrische Deutung der Vektoraddition	22
1.3.2	Vektoraddition und Translationen	23
1.4	Multiplikation von Vektoren mit Skalaren	24
1.4.1	Vektorraum allgemein	26
1.4.2	Ein Funktionenraum als Vektorraum	26
1.4.3	Linearkombinationen	26
1.4.4	Spezielle Linearkombinationen	28
1.5	Lineare Unabhängigkeit, Basis	29
1.5.1	Der 3-dimensionale Fall	29
1.5.2	Der n-dimensionale Fall	30
1.5.3	Der ebene Fall	32
1.6	Länge von Vektoren	34
1.6.1	Abstand zwischen zwei Punkten	34
1.6.2	Normierung auf Länge Eins	34
1.7	Geraden	35
1.7.1	Koordinatenform von Geraden in der euklidischen Ebene	37
1.7.2	Kollinearität von Punkten	38
1.7.3	Schnitt zweier Geraden	38
1.7.4	Wie löst man ein (ebenes) lineares Gleichungssystem?	39
1.8	Ebenen im dreidimensionalen Raum	39
1.8.1	Parallele Ebenen	41

1.8.2	Komplanarität	43
1.8.3	Schnitt von Ebene und Gerade	44
1.8.4	Schnittgerade zweier Ebenen	45
1.8.5	Koordinatenform von Ebenen	46
2	Skalarprodukt, Winkel, Orthogonalität	49
2.1	Einführung	49
2.2	Skalarprodukt von zwei Vektoren und Winkel zwischen Vektoren	50
2.2.1	Rechenregeln für das Skalarprodukt	53
2.2.2	Länge eines Vektors und Skalarprodukt	54
2.2.3	Formel für die Fläche eines Parallelogramms	55
2.2.4	Korrelationen und Skalarprodukt	56
2.2.5	Ein physikalisches Beispiel	57
2.3	Orthogonalität	58
2.3.1	Satz von Pythagoras	58
2.3.2	Parallelogrammgleichung	60
2.3.3	Orthogonalbasis	60
2.4	Geraden: Hessesche Normalform	62
2.4.1	Lot auf Geraden	62
2.4.2	Geradendarstellung mit Hilfe des Skalarprodukts; Hessesche Normalform	62
2.4.3	Umrechnungen von Koordinaten- in Parameterform und umgekehrt	63
2.5	Ebenen: Hessesche Normalform	64
2.5.1	Umrechnungen von Koordinaten- in Parameterform und umgekehrt	65
2.6	Vektorprodukt	65
2.6.1	Berechnung des Vektorprodukts	67
2.6.2	Vektorprodukt und Determinanten von 2x2-Matrizen	68
2.6.3	Das Volumen eines Parallelepipeds	68
2.6.4	Determinante einer 3x3-Matrix	69
2.6.5	Internet	70
3	Lineare Abbildungen	71
3.1	Einführung	71
3.2	Lineare Gleichungssysteme	72
3.2.1	Bisherige Beispiele	72
3.2.2	Formulierung eines allgemeinen linearen Gleichungssystems	74
3.2.3	Anwendungen	76
3.2.4	Elementare Zeilenumformungen	77
3.2.5	Gauß'sches Eliminationsverfahren	77
3.3	Lineare Abbildungen und Matrizen	79
3.3.1	Linearität im Nahrungsmittel-Nährstoff-Beispiel	80
3.3.2	Lineare Gleichungssysteme und Lineare Abbildungen	80

3.3.3	Linearität	81
3.3.4	Geometrische Transformationen und lineare Abbildungen	82
3.3.5	Ausnutzung der Linearität	85
3.3.6	Internet	86
3.3.7	Verkettung von linearen Abbildungen – Multiplikation von Matrizen	86
3.4	Zentralprojektion	90
3.5	Parallelprojektion	92

Teil II

Lineare Algebra und analytische Geometrie

Kapitel 1

Der Anschauungsraum und seine n-dimensionale Form

1.1 Einführung

In diesem Kapitel wollen wir den euklidischen Raum kennenlernen, der mit unserer Anschauung übereinzustimmen scheint. Deshalb bezeichnet man ihn auch als *Anschauungsraum* im 3-dimensionalen und mit *Anschauungsebene* im 2-dimensionalen Fall.

Die *Punkte* der Anschauungsebene kann man durch Zahlenpaare, also durch Elemente des \mathbb{R}^2 , die des Anschauungsraums durch Zahlentripel, also durch Elemente des \mathbb{R}^3 beschreiben, wenn ein Koordinatensystem, etwa ein kartesisches Koordinatensystem gegeben ist. \mathbb{R}^2 und \mathbb{R}^3 , allgemein der \mathbb{R}^n , heißen *euklidische Räume* und sind die einfachsten und wichtigsten Beispiele von Vektorräumen, mit denen sich die *Lineare Algebra* beschäftigt. EUKLID, siehe Abb. 1.1 lebte etwa 300 v.Chr. und ist der Begründer der klassischen *axiomatischen Geometrie*. Euklidische Räume zeichnen sich durch *Skalarprodukte* aus, mit deren Hilfe man *Winkel* und *Längen* messen kann. Schon der Name *kartesisches* Koordinatensystem besagt, dass die Koordinatenachsen senkrecht aufeinander stehen, was nur einen Sinn gibt, wenn man etwas über Winkel weiß.

Von der Schule kennen Sie die Anschauungsebene in Form eines Blattes Papier oder der Tafel, die zwei mit einer Skala versehene Koordinatenachsen enthält. Diese Koordinatenachsen erlau-



Abbildung 1.1: Euklid von Alexandria, ca. 300 v.Chr.

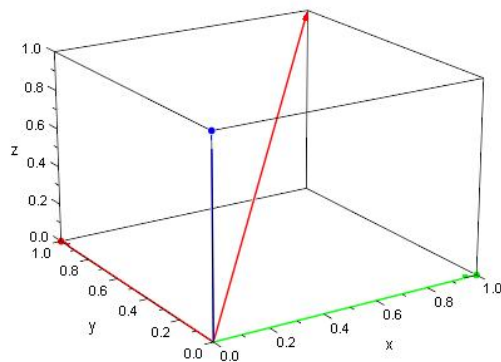


Abbildung 1.2: Der 3-dimensionale Anschauungsraum

ben eine genaue Lokalisierung eines Punktes in Gestalt eines geordneten Paares $P := (x, y)$ des kartesischen Produkts $\mathbb{R}^2 = \mathbb{R} \times \mathbb{R}$ durch die *Koordinaten* (oder *Komponenten*) von P . Sie führen zur Beschreibung von *Geraden*, zum Begriff *Abstand* zwischen Punkten, u.v.m.

Ganz analog — aber doch schon ein wenig unhandlicher — werden Punkte des Anschauungsraums durch Zahlentripel $P = (x, y, z) \in \mathbb{R}^3$ lokalisiert, wobei man sich eine dritte Koordinatenachse, die z.B. senkrecht aus der Tafelenebene herausragt, vorstellen muss, s. Abb. 1.2. Dies führt z.B. zur Beschreibung von *Ebenen*.

Diese kartesischen Koordinatensysteme mit senkrecht aufeinander stehenden Koordinatenachsen sind bestens geeignet, anschauliche Geometrie im „Kleinen“ zu betreiben, z.B. bei der Beschreibung von gewissen technischen „Körpern“, wie z.B. irgend ein Werkzeug. Oder auch in der Architektur — denken Sie an 3D-Ansichten eines Gebäudes. Sie sind jedoch ungeeignet, um z.B. den Ort eines Schiffes oder eines Flugzeugs oder gar eines Himmelskörpers festzulegen. Hierzu sind andere Koordinatensysteme, etwa *Kugelkoordinaten* mit *Breite* und *Länge* nötig. Gänzlich unzulänglich ist die Geometrie des euklidischen Raumes, wenn es um Raum und Zeit in unserem Kosmos geht (*Relativitätstheorie*).

Koordinatensysteme (häufig auch nicht-kartesische!) dienen also zur *Ortsbestimmung*. Sie kennen bereits Koordinaten von Stadtplänen („Planquadrate“ dienen der groben Lokalisierung) und Atlanten. Aber auch auf „krummen“ Autobahnen gibt es Koordinaten in Form von „Kilometersteinen“ — als eine Realisierung eines eindimensionalen Raums, der mathematisch durch eine *Zahlengerade* gegeben ist. Etwas komplizierter ist die Positionsangabe von Himmelskörpern, für die es verschiedene Koordinatensysteme gibt (das natürlichste ist das Horizontalsystem). Im Internet gibt es unter [Astronomische Koordinaten](#) serh schöne Erläuterungen der verschiedenen Koordinatensysteme.

Die *analytische Geometrie* ist eng mit der *linearen Algebra* verbunden. Sie befasst sich mit der Analyse geometrischer Fragen mit Hilfe der *Vektorrechnung*. Als Beispiel nenne ich die Berechnung des Flächeninhalts eines Parallelogramms mit Hilfe der Koordinaten der Eckpunkte (hier gibt es einen schönen Zusammenhang mit der *Determinante* einer *Matrix*).

Die eben angesprochene Vektorrechnung gibt aber häufig auch Sinn bei nicht-geometrischen

Problemen. So haben wir es in der Statistik bei einer Stichprobe vom Umfang n mit einem n -Tupel (x_1, x_2, \dots, x_n) von Daten in Form reeller Zahlen zu tun, die analysiert werden können — ohne einen direkten Bezug zur Geometrie.

Zu Vektorräumen mit einer *linearen Struktur* wird der \mathbb{R}^n durch zwei elementare Operationen: die *Vektor-Addition* und die *Skalarmultiplikation* — die Multiplikation eines Vektors mit einer reellen Zahl¹. *Punkte* und *Vektoren* sind mathematisch nicht zu unterscheiden — es handelt sich um Zahlenpaare oder Zahlentripel (oder um n -Tupel im \mathbb{R}^n , in der Datenverarbeitung einfach nur um *Listen* reeller Zahlen).

Dennoch haben Punkte und Vektoren *geometrisch* eine unterschiedliche Bedeutung. Da die Lineare Algebra aber nicht nur auf geometrische Anwendungen aus ist, gibt es außerhalb der Geometrie keinen Grund, diese beiden Objekte zu unterscheiden. Ich werde dies aber dennoch immer dann tun, wenn die Geometrie zur Motivation der Konzepte unerlässlich ist. Es ist schon bemerkenswert, dass sich die von der Geometrie herkommenden Konzepte (wie *Winkel* zwischen Vektoren) auch in nicht-geometrischen Bereichen — wie z.B. in der Statistik mit Korrelationen — als ausgesprochen nützlich herausgestellt haben.

Meist werde ich Vektoren und Punkte mit fetten Buchstaben wie \mathbf{a} , \mathbf{b} , \mathbf{x} , \mathbf{P} , ... bezeichnen. Wenn jedoch in einer geometrischen Situation sehr wohl zwischen Punkten und Vektoren unterschieden werden soll, werde ich nur die Vektoren „einfetten“, nicht aber die Punkte.

Die Lineare Algebra befasst sich aber nicht nur mit Vektorrechnung, sondern auch mit gewissen, häufig geometrisch deutbaren *Abbildungen* eines euklidischen Raums in sich oder in einen anderen. Hierdurch können z.B. *Bewegungen* eines starren Körpers beschrieben werden. Beispiele für solche Abbildungen sind *Translationen (Verschiebungen)*, *Drehungen*, *Spiegelungen*, *Scherungen*, *Projektionen*, Hierbei kommt der Begriff der *Matrix* in Verbindung mit *linearen Abbildungen* zum Zuge. Matrizen sind eigentlich nichts anderes als rechteckige Tabellen von Zahlen — für jeden Excel-Anwender ein Begriff. Solche Tabellen verwandeln sich in der analytischen Geometrie zu linearen Abbildungen des Raums. Z.B. wird die simple Tabelle

$$\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$$

zu einer Drehung um 90 Grad, die Zuordnungsvorschrift lautet $(x, y) \mapsto (-y, x)$. Es ist ganz wichtig, dass Sie lernen sich von dem Funktionsbegriff einer reellen Funktion, wie Sie sie aus der Schule (Stichwort: „Funktionsgleichungen“) kennen, zu lösen und ihn z.B. auf solche Abbildungen von \mathbb{R}^2 nach \mathbb{R}^2 übertragen zu können.

Es wird vier Kapitel geben: in dem ersten geht es um die Vektorrechnung in *Vektorräumen* (die hier fast immer der \mathbb{R}^n , speziell für $n = 2, 3$ ist). Während wir in diesem ersten Teil noch nichts über *Winkel* zwischen Vektoren sagen², werden wir im zweiten Teil (Kap. 2) ein *Skalarprodukt* auf dem \mathbb{R}^n einführen, das zur Berechnung von Winkeln führt. Im dritten Teil (Kap. 3) geht

¹Nicht zu verwechseln mit dem Skalarprodukt zwischen zwei Vektoren.

²wenn man von der Erwähnung von „kartesischen Koordinatensystemen“ absieht.

es um (lineare) Abbildungen, die z.B. in der *Computergrafik* eine Rolle spielen. Dort will man gewisse *Objekte*, das sind gewisse Punktmengen manipulieren, z.B. verschieben oder um eine Achse drehen oder eine Verzerrung vornehmen. Mathematisch sind dies alles Abbildungen des euklidischen Raums in sich, im einfachsten Fall handelt es sich um *lineare Abbildungen* oder auch um *Bewegungen* – das sind solche Abbildungen, die Längen und Winkel sowie die Orientierung unverändert lassen: Z.B. wird aus einem Dreieck ein deckungsgleiches Dreieck, dessen Eckpunkte in der gleichen Orientierung durchlaufen werden wie vor der Abbildung. Die *Lineare Algebra* befasst sich mit den Eigenschaften solcher *linearen Abbildungen*.

Ein wesentliches Element der 3D-Computergrafik ist die Beschreibung von 3D-Objekten durch *Projektionen* als 2D-Objekte. Hiermit hängen die verschiedenen *Perspektiven* zusammen, die mathematisch als Abbildungen $F : \mathbb{R}^3 \rightarrow \mathbb{R}^2$ beschrieben werden können. In der Geographie hängt diese Frage damit zusammen, wie der Globus auf einzelnen ebenen Seiten dargestellt werden kann.

Als letztes sei vermerkt, dass die Vektorrechnung eine nicht wegzudenkende Bedeutung in der Physik hat. Viele physikalischen Begriffe wie *Kraft*, *Impuls*, *Geschwindigkeit*, *Ort*, *Drehmoment*, *Stromstärke*, *Magnetfeld*, korrespondieren mit Vektoren. Die Mathematik hat hier ganz wesentliche Anregungen erhalten und gegeben. Ich werde jedoch diese physikalischen Bezüge nicht in den Vordergrund stellen, weil ich glaube, dass viele von Ihnen dies eher als Erschwernis auffassen würden.

Im letzten Kapitel wird die abstrakte Lineare Algebra mit ihren Begriffen *Vektorraum*, *Dimension*, *Basis*, *linearen Abbildungen* vorgestellt.

Internet: [Vector calculator](#) (Uni Nizza - Frankreich, auf englisch), [Einführung in Vektoren](#) (Mathe Online Wien) [Vektorrechnung](#) (Online-Skript Uni-Stuttgart)

Die hier und im nächsten Kapitel präsentierten Inhalte orientieren sich zum Teil an den 1979 erschienenen Texten zum *Fernstudium im Medienverband (FIM)* und an den Internet-Texten von *Mathe Online (Wien)*. Die Zeichnungen stammen meist aus diesen beiden Quellen.

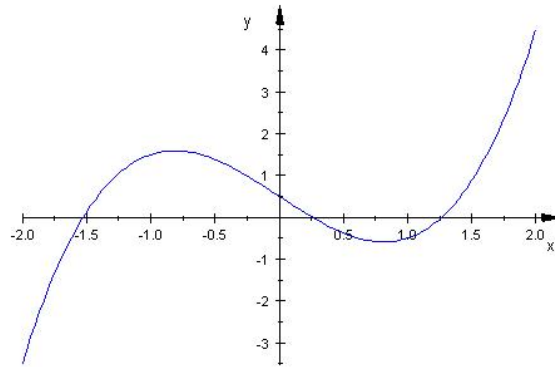
1.1.1 Rückblick auf Mathe I

Das *kartesische Produkt* $\mathbb{R}^n := \mathbb{R} \times \mathbb{R} \times \cdots \times \mathbb{R}$ (n -mal) ist uns vertraut³. Für $n = 2$ haben wir die Anschauungsebene und für $n = 3$ den Anschauungsraum. Wir müssen uns nur immer wieder klar machen, welche Informationen uns die einzelnen Komponenten eines Elementes („Punktes“) $x := (x_1, x_2, \dots, x_n) \in \mathbb{R}^n$ liefern.

Wir haben schon Methoden kennengelernt, gewisse Punktmengen des \mathbb{R}^2 zu beschreiben. So ist

$$G(f) := \{(x, y) \in \mathbb{R}^2 : y = f(x)\}$$

³In welchem Sinne können Sie \mathbb{R}^n als eine Potenz mit Exponentem n auffassen?

Abbildung 1.3: Funktionsgraph von $f(x) := x^3 - 2x + 0.5$, $-2 \leq x \leq 2$

der Graph einer reellen Funktion $f : \mathbb{R} \rightarrow \mathbb{R}$. Allgemeiner ist jede Relation auf \mathbb{R} ja eine Teilmenge von \mathbb{R}^2 . Haben Sie schon einmal darüber nachgedacht, wie die zur Ordnungsrelation

$$(x, y) \in R : \iff x \leq y$$

gehörende Relation R auf \mathbb{R} gezeichnet werden kann?

Können Sie sich an die Diskussion in einer Vorlesung erinnern, wie die Menge

$$K := \{(x, y) \in \mathbb{R}^2 : |x| + |y| \leq 1\}$$

aussieht? Oder an die Ferienaufgaben mit

$$M_1 := \{(x, y) \in \mathbb{R}^2 : |x - 1| + |y - 2| \leq 4\},$$

$$M_2 := \{(x, y) \in \mathbb{R}^2 : |x - 1| - |y - 2| \leq 4\},$$

$$M_3 := \{(x, y) \in \mathbb{R}^2 : ||x - 1| - |y - 2|| \leq 4\}?$$

Wenn $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ eine Funktion in zwei (reellen) Variablen ist, so beschreibt die *Funktionsgleichung* $f(x, y) = 0$ eine Lösungsmenge

$$A := \{(x, y) \in \mathbb{R}^2 : f(x, y) = 0\},$$

die für eine in x und y quadratische Funktion $(x, y) \mapsto f(x, y) = a_0 + a_1x + a_2y + a_3x + a_4x^2 + a_5y^2$ ein *Kegelschnitt* ist, z.B. eine Ellipse oder eine Hyperbel oder eine Parabel. Auch Kreise oder Geraden oder gar Punkte können darunter fallen (s. Kap. 5.3 und Abb. 1.5 nach meinem Lehrer L. COLLATZ (1910 1990)).

Auf die Lineare Algebra mit ihren linearen Abbildungen zielte schon die Aufgabe 44 aus Mathe I. So ist die Spiegelung S_g an der Geraden $g := \{(x, y) : y = 0\}$ offensichtlich durch $S_g : (x, y) \mapsto (x, -y)$ definiert. Jede Spiegelung S_g an einer Geraden g ist selbstinvers, d.h. S_g ist (wie auch

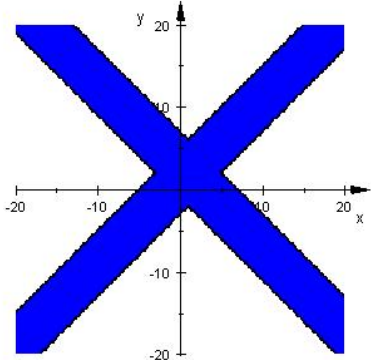


Abbildung 1.4: Die Menge M_3

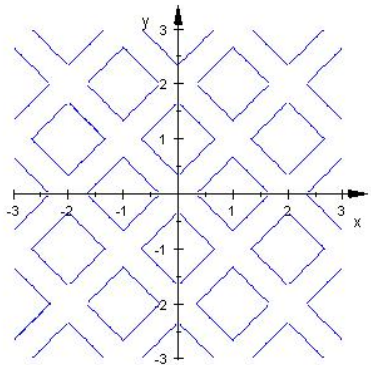


Abbildung 1.5: L. COLLATZ: $f(x, y) = |||x| - 1| - 1| - |||y| - 1| - 1|| - 1/3$

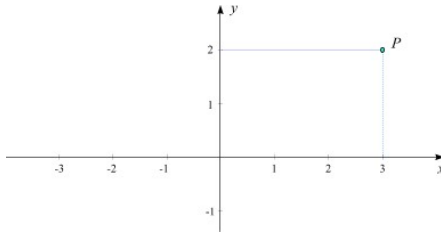


Abbildung 1.6: Punkt in der Ebene

jede andere Spiegelung) die Umkehrabbildung von sich selbst. Es wird sich herausstellen, dass S_g ein Beispiel einer *linearen Abbildung* der Ebene in sich ist, wenn g durch den Ursprung geht. Wie auch eine Drehung um 180 Grad, die durch $(x, y) \mapsto (-x, -y)$ definiert ist.

In der Gaußschen Zahlenebene haben Sie sowohl kartesische Koordinaten ($z = x + iy$) als auch Polarkoordinaten r und ϕ ($z = r(\cos \phi + i \sin \phi)$) kennengelernt. Insofern kann man jede komplexe Zahl z als einen Punkt (x, y) oder (r, ϕ) der Ebene auffassen. Funktionen $f : \mathbb{C} \rightarrow \mathbb{C}$ können somit auch als Funktionen $F : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ geschrieben werden. Z.B. hat $f(z) := \bar{z}$ die Darstellung $F(x, y) := (x, -y)$, wobei F wieder als Spiegelung an der x -Achse gedeutet werden kann.

1.2 Vektoren

1.2.1 Punkte in der Anschauungsebene und im Anschauungsraum

Wir kennen bereits das kartesische Produkt $\mathbb{R}^2 := \mathbb{R} \times \mathbb{R}$ und die Deutung seiner Elemente als *Punkte* einer durch ein kartesisches Koordinatensystem⁴ markierten *Ebene*. Die Punkte $\mathbf{x} \in \mathbb{R}^2$ haben **Koordinaten**⁵ (oder **Komponenten**) x_1 und x_2 : $\mathbf{x} = (x_1, x_2)$. In Abb. 1.6 sind die Achsen mit x und y beschriftet, es wird der Punkt $\mathbf{P} = (3, 2)$ dargestellt. Bei Dezimalzahlen greife ich wegen ihrer Kommadarstellung auf eine andere Notation zurück: Das Komma, das im geordneten Paar die Komponenten trennen soll, wird durch ein Semikolon ersetzt, z.B. $\mathbf{x} = (1, 23; 3, 45)$, zuweilen auch durch einen senkrechten Strich: $\mathbf{x} = (1, 23|3, 45)$.

Den (dreidimensionalen) Raum (3D-Raum) beschreiben wir durch das kartesische Produkt \mathbb{R}^3 . Dessen Punkte $\mathbf{x} \in \mathbb{R}^3$ haben drei Koordinaten, $\mathbf{x} = (x_1, x_2, x_3)$, meist spricht man auch von den x -, y - und z -Koordinaten. In \mathbb{R}^3 gibt es drei Koordinatenachsen, die sich nicht so einfach auf eine Tafel- oder Zeichenebene zeichnen lassen. Es gibt verschiedene „Perspektiven“, von denen einige in den Abb. 1.7–1.10 zu sehen sind. Die Achsen sind mit x , y und z beschriftet⁶.

⁴nach Rene Descartes, 1596 - 1650

⁵Aus „didaktischen“ Gründen spreche ich nicht von x - und y -Koordinaten, sondern von erster und zweiter Koordinate. Die Achsen des Koordinatensystems werden als x_1 - und x_2 -Achse oder als Abszisse und Ordinate bezeichnet. Der Buchstabe y ist für andere Punkte reserviert.

⁶Wie in der Einleitung bemerkt, kann es auch Zahlentripel geben, die eine geometrische Deutung nicht

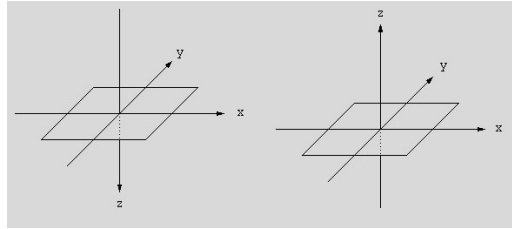


Abbildung 1.7: 3D-Koordinatensystem

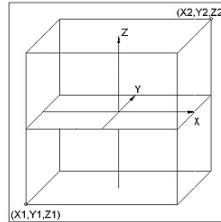


Abbildung 1.8: 3D-Koordinatensystem

Wir ziehen bei einer geometrischen Veranschaulichung vor, die Höhe durch $x_3 = z$, die Breite durch $x_1 = x$ und die Tiefe durch $x_2 = y$ zu messen, siehe Abb. 1.10. Dies entspricht der Situation an einer Tafel oder am Bildschirm. Hat man jedoch eine Zeichenebene auf einem Tisch, wird man ausgehend vom üblichen ebenen Koordinatensystem mit $x_1 = x$ -Achse und $x_2 = y$ -Achse sich die dritte Achse senkrecht nach oben gerichtet denken.

Wir werden erst im nächsten Kapitel mit Hilfe eines *Skalarprodukts* den Begriff „senkrecht“ erfassen, insofern spielen solche kartesischen Koordinatensysteme noch eine untergeordnete Rolle.

Einen guten Einstieg in das interaktive Umgehen in einem 3D-Koordinatensystem gibt das [Applet: Räumliche Koordinaten](#) (Mathe Online Wien). Hier können Sie das Koordinatensystem um die Koordinatenachsen drehen und Sie können zoomen.

Aber nicht nur einzelne Punkte, sondern auch *Punktmengen* sind Gegenstand der Vektorrech-

notwendig machen. Z.B., wenn man jeder Person ein Zahlentripel bestehend aus Gewicht, Länge und Alter zuordnet.

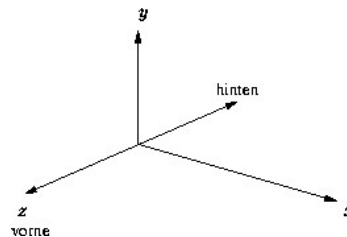


Abbildung 1.9: 3D-Koordinatensystem

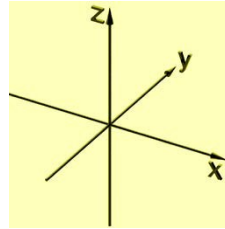


Abbildung 1.10: 3D-Koordinatensystem

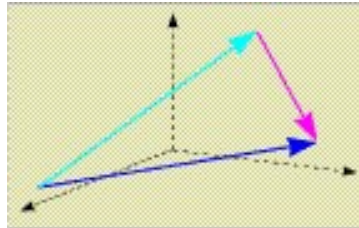


Abbildung 1.11: Vektoren als Pfeile

nung. So kann eine Kreisscheibe mit Mittelpunkt $(1, 1)$ und Radius 2 durch $\{(x, y) \in \mathbb{R}^2 : (x - 1)^2 + (y - 1)^2 \leq 4\}$, der zugehörige Kreis durch $\{(x, y) \in \mathbb{R}^2 : (x - 1)^2 + (y - 1)^2 = 4\}$ beschrieben werden. Die Punktmenge $\{(x, y) \in \mathbb{R}^2 : y + x = 4\}$ stellt eine Gerade mit Steigung -1 und Achsenabschnitt 4 dar.

1.2.2 Vektoren als „Pfeile“

Wir haben bisher von *Punkten* in der Ebene und im Raum gesprochen. Oder einfach von n -Tupeln. In der *Linearen Algebra* zieht man es jedoch vor, von *Vektoren* zu sprechen, weil mit diesem Konzept die (Vektor-) Addition leichter motiviert werden kann. Dennoch sollte man in der *analytischen Geometrie* zuweilen sehr wohl zwischen Punkten und Vektoren unterscheiden. Letztere werden gerne durch „Pfeile“ markiert, die einen Schaft und eine Spitze in Form von Punkten haben.

Es gibt im Wesentlichen drei verschiedene, aber miteinander verwandte geometrische Veranschaulichungen eines Vektors als Pfeil:

- **Verbindungsvektor**

Sind zwei Punkte P und Q gegeben, so nennt man den Pfeil von P nach Q einen **Verbindungsvektor**, bezeichnet durch $\mathbf{a} := \vec{PQ}$, s. Abb. 1.12. Schon jetzt sei vermerkt, dass man die Komponenten von \mathbf{a} durch die „Spitze minus Schaft-Regel“ berechnet (siehe weiter unten). Hieraus ergibt sich auch, dass es andere Punktepaare P, Q gibt, die zu demselben Verbindungsvektor führen. Ein Pfeil ist also nur ein Repräsentant eines Vektors, man kann ihn beliebig (parallel) verschieben, s. Abb. 1.13.

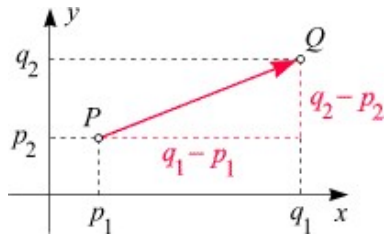


Abbildung 1.12: Verbindungsvektor

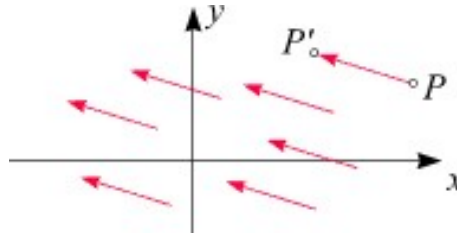


Abbildung 1.13: Verschiebung von Vektoren

- **Verschiebungsvektor**

Wir knüpfen an die letzte Bemerkung an, die eine weitere (die „mathematischste“) Deutung eines Vektors als **Verschiebungsvektor** zulässt. Zu jedem Vektor \mathbf{a} gibt es eine Abbildung (auch **Translation** genannt) $T_{\mathbf{a}}$, die Punkte so verschiebt⁷, dass \mathbf{a} gerade der Verbindungsvektor von Urbild- und Bildpunkt ist, s. Abb. 1.13. Will man „Formeln“ für $T_{\mathbf{a}}(P)$ angeben, kommt man zwangsläufig auf die Vektoraddition. s. Kap. 1.3.2.

- **Ortsvektor**

⁷Erinnern Sie sich, wie man mit Hilfe von Lineal und Zeichendreieck eine Translation konstruiert hat?

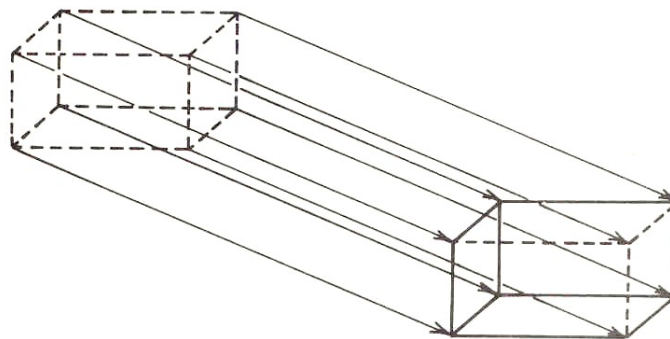


Abbildung 1.14: Parallelverschiebung eines quaderförmigen Körpers

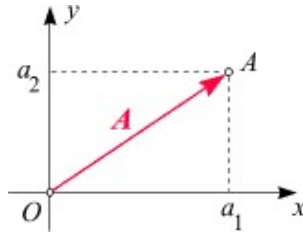


Abbildung 1.15: Ortsvektoren

Ist ein Punkt A gegeben, so ist dessen **Ortsvektor** \mathbf{A} gerade der Verbindungsvektor zwischen Ursprung und Punkt P , s. Abb. 1.15.

„**Spitze minus Schaft-Regel**“: Sind die Koordinaten des Anfangspunktes P und des Endpunktes Q des Verbindungsvektors $\mathbf{a} = \vec{PQ}$ gegeben, so können wir die Komponenten von \mathbf{a} ganz einfach berechnen — man muss sich den Punkt P nur in den Ursprung verschoben denken: Sind in der Ebene die Punkte $P = (p_1, p_2)$ und $Q = (q_1, q_2)$ gegeben, so gilt (s. Abb. 1.12)

$$\mathbf{a} = (q_1 - p_1, q_2 - p_2).$$

Entsprechend gilt im Raum für die Punkte $P = (p_1, p_2, p_3)$ und $Q = (q_1, q_2, q_3)$:

$$\mathbf{a} = (q_1 - p_1, q_2 - p_2, q_3 - p_3).$$

Im nächsten Abschnitt, wenn wir lernen, dass man Vektoren addieren (und subtrahieren) kann, werden wir diese Regel als eine Beziehung zwischen Orts- und Verbindungsvektor erkennen: Es gilt $\mathbf{a} = \vec{PQ} = \mathbf{Q} - \mathbf{P}$, wenn \mathbf{P} und \mathbf{Q} die zu den Punkten P und Q zugehörigen Ortsvektoren sind.

1.2.3 Vektoren im n-dimensionalen Raum

Wir haben Vektoren (bzw. Punkte) bisher in der Form $\mathbf{x} = (x_1, x_2)$ oder $\mathbf{x} = (x_1, x_2, x_3)$ dargestellt. Im ersten Fall gibt es zwei, im zweiten drei Komponenten. Natürlich können wir uns jetzt auch den *n-dimensionalen euklidischen Raum* \mathbb{R}^n vorstellen, dessen Vektoren die Form $\mathbf{x} = (x_1, x_2, \dots, x_n)$ mit n Komponenten haben. Die Deutung von solchen Vektoren mit Hilfe von Pfeilen ist hier in keiner Weise hilfreich!

Diese Darstellungsform von Vektoren nennt man auch **Zeilenvektoren** — im Gegensatz zu **Spaltenvektoren**⁸

⁸Es wird sich zeigen, dass man mit Spaltenvektoren besser rechnen kann. Zeilenvektoren sind jedoch bequemer und nehmen nicht so viel Platz ein.

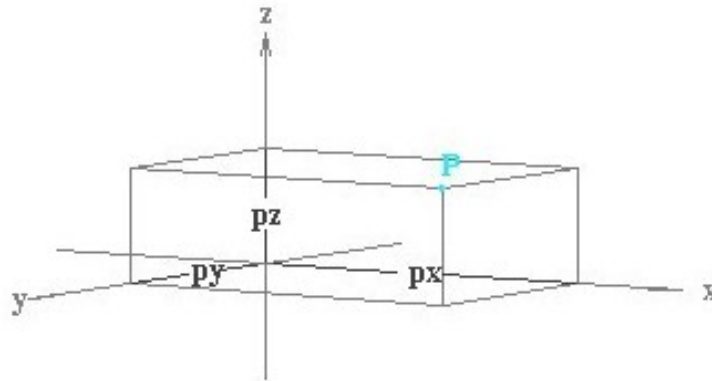


Abbildung 1.16: 3D-Koordinatensystem Applet Mathe Online

$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$

Auch $\mathbb{R} = \mathbb{R}^1$ ist ein Raum von (1-dimensionalen) Vektoren $\mathbf{x} = (x_1)$. Wir unterscheiden aber nicht zwischen x_1 und (x_1) . Insofern können wir auch von reellen Zahlen als 1D-Vektoren sprechen, die wir auf dem Zahlenstrahl veranschaulichen⁹.

Aufgabe: Welche der Achsenlagen in den Abb. 1.7–1.10 können Sie durch Drehen ineinander überführen¹⁰? Die Ausgangslage der Achsen im [Applet: Räumliche Koordinaten](#) (Mathe Online Wien) sieht man in Abb. 1.16.

1.2.4 Vektoren in der Statistik

Vektoren haben nicht immer mit *geometrischen Objekten* zu tun. So werden in der Statistik mittels einer *Stichprobe vom Umfang n* gewisse Quantitäten (Zahlen) wie Geburtsgewicht, Länge, Alter, Intelligenzquotient gemessen und zu einem *Stichprobenvektor* $\mathbf{x} = (x_1, x_2, \dots, x_n)$ zusammengestellt, wobei x_j die gemessene Größe der j -ten Person in der Stichprobe ist. Auch hier gibt es naheliegende Abbildungen, z.B. diejenige, die dem Vektor $\mathbf{x} \in \mathbb{R}^n$ in einen *Mittelwert* zurordnet, also $\mathbf{x} \mapsto \bar{x} := \frac{\sum_{j=1}^n x_j}{n}$ oder diejenige Abbildung, die die *empirische Varianz*

⁹Meines Wissens wird die Pfeilschreibweise auch auf dem Zahlenstrahl benutzt, um Addition und Subtraktion von Zahlen bzw. die Operationen $x \mapsto x + a$ und $x \mapsto x - a$ zu veranschaulichen.

¹⁰Diese Frage hat etwas mit der Orientierung des Koordinatensystems zu tun. Oder auch mit der „Rechte-Hand-Regel“, besser der Dreifinger-Regel bestehend aus Daumen, Zeige- und Mittelfinger. Mehr hierzu im nächsten Kapitel.

$\sigma^2 := \frac{\sum_{j=1}^n (x_j - \bar{x})^2}{n-1}$ ermittelt¹¹.

1.3 Vektoraddition

Wir definieren eine **Addition von Vektoren** des \mathbb{R}^2 bzw. des \mathbb{R}^3 durch

$$(x_1, x_2) + (y_1, y_2) := (x_1 + y_1, x_2 + y_2)$$

bzw. durch

Definition 1.1.

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

Diese Addition als binäre Verknüpfung macht $(\mathbb{R}^2, +)$ bzw. $(\mathbb{R}^3, +)$ zu einer *abelschen Gruppe*, s. Def. I.8.9. Eine Verallgemeinerung auf den \mathbb{R}^n ist offensichtlich:

Definition 1.2. Die Addition zweier Vektoren des \mathbb{R}^n wird komponentenweise definiert:

$$(x_1, x_2, \dots, x_n) + (y_1, y_2, \dots, y_n) := (x_1 + y_1, x_2 + y_2, \dots, x_n + y_n).$$

Beachten Sie, dass das neutrale Element bzgl. der Addition der **Nullvektor** $\mathbf{o} = (0, 0, \dots, 0)$ ist. Zu $\mathbf{u} = (u_1, u_2, \dots, u_n)$ invers ist $(-u_1, -u_2, \dots, -u_n)$.

Damit gelten die folgenden Rechenregeln einer abelschen Gruppe:

- (Kommutativität) $\mathbf{v}_1 + \mathbf{v}_2 = \mathbf{v}_2 + \mathbf{v}_1$
- (Assoziativität) $(\mathbf{v}_1 + \mathbf{v}_2) + \mathbf{v}_3 = \mathbf{v}_1 + (\mathbf{v}_2 + \mathbf{v}_3)$
- (neutrales Element) $\mathbf{v} + \mathbf{o} = \mathbf{v}$
- (inverse Elemente) Zu jedem \mathbf{v} gibt es ein inverses Element \mathbf{w} mit $\mathbf{v} + \mathbf{w} = \mathbf{o}$.

Bezeichnet man das additive Inverse von \mathbf{v} mit $\mathbf{w} := -\mathbf{v}$, so kann man eine Subtraktion von Vektoren durch $\mathbf{v}_1 - \mathbf{v}_2 = \mathbf{v}_2 + (-\mathbf{v}_1)$ definieren. Jetzt ergibt sich die Spitze-minus-Schaft-Regel $\vec{PQ} = Q - P$.

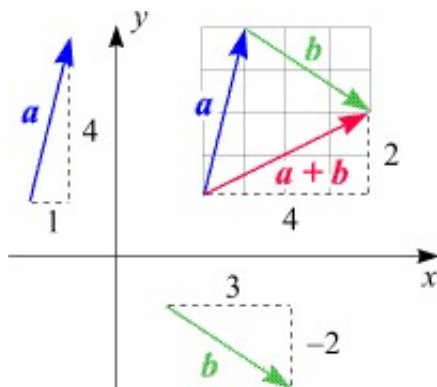
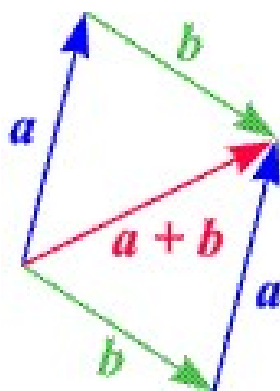


Abbildung 1.17: Addition von ebenen Vektoren

Abbildung 1.18: $\mathbf{a} + \mathbf{b} = \mathbf{b} + \mathbf{a}$

1.3.1 Geometrische Deutung der Vektoraddition

Algebraisch gibt es an der Definition 1.2 nichts rätselhaftes. Wir kennen diese Addition auch schon von der komplexen Zahlenebene, insofern ist sie etwas vertraut.

Wenn man auf die „Pfeildeutung“ von Vektoren, s. Abb. 1.17, zurückgreift, ergibt sich ein neuer Aspekt:

Erst zeichnen wir \mathbf{a} und setzen daraufhin den Schaft des Vektors \mathbf{b} an die Spitze von \mathbf{a} an. Der neue Vektor $\mathbf{a} + \mathbf{b}$ hat dann den Schaft von \mathbf{a} als Schaft und die Spitze von \mathbf{b} als Spitze.

Wenn man die Reihenfolge in dieser „Pfeilkonstruktion“ vertauscht, also erst \mathbf{b} und dann \mathbf{a} zeichnet, erhält man dasselbe Ergebnis¹², insgesamt entsteht ein *Parallelogramm*, siehe Abb. 1.18.

Das entstehende Parallelogramm wird auch als *Kräfteparallelogramm* bezeichnet, siehe Abb. 1.19 mit den (*Kraft-*) Vektoren F_1 und F_2 . Die Summe der beiden in einem gemeinsamen Punkt angreifenden Kräfte ist dann die resultierende Kraft. In dem Schaft des Vektors

¹¹Der Nenner $n - 1$ im Ausdruck für σ^2 ist *kein* Druckfehler.

¹²Eine Konsequenz von $\mathbf{a} + \mathbf{b} = \mathbf{b} + \mathbf{a}$.

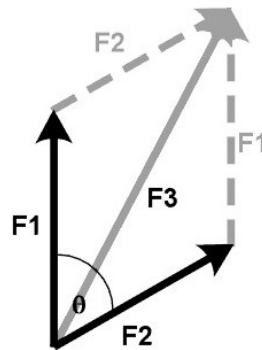


Abbildung 1.19: Addition von ebenen Vektoren

greift die Kraft an, der Pfeil gibt die Richtung der Kraft an, seine Länge die Stärke der Kraft. Zwei Kräfte, die gleiche Länge, aber entgegengesetzte Richtung haben, heben sich auf, Ihre Summe ist der Nullvektor \mathbf{o} .

1.3.2 Vektoraddition und Translationen

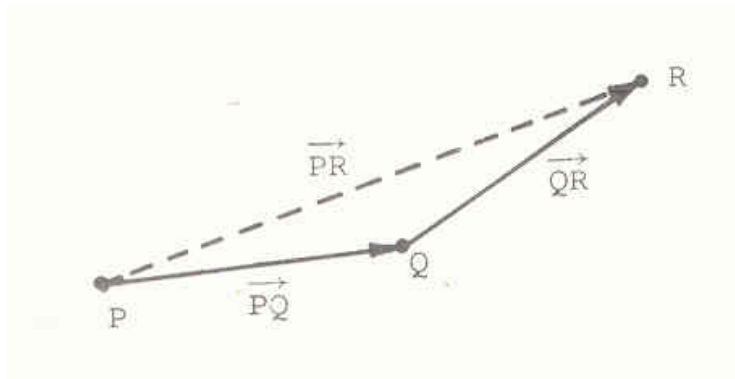
Man betrachtet ganz spezielle Abbildungen $T : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ bzw. $T : \mathbb{R}^3 \rightarrow \mathbb{R}^3$, allgemein $T : \mathbb{R}^n \rightarrow \mathbb{R}^n$, die sogenannten *Translationen*, zu deutsch (*Parallel-*) *Verschiebungen*. Jeder Punkt wird dabei in eine bestimmte Richtung um eine bestimmte Länge verschoben – mit einem *Verschiebungsvektor* $\mathbf{a} := (a_1, a_2, \dots, a_n)$. Die Translation ist durch $T : \mathbf{x} \mapsto \mathbf{x} + \mathbf{a}$ definiert. Um die eindeutige Beziehung zwischen *Vektor* \mathbf{a} und *Translation* T zu erfassen, schreiben wir $T_{\mathbf{a}}$ an Stelle von T .

Jetzt machen wir ein Gedankenspiel: Wir betrachten die Abbildung, die jedem Vektor \mathbf{a} die *zugehörige* Translation $T_{\mathbf{a}}$ zuordnet. Diese Abbildung hat als Definitionsbereich die Menge aller Vektoren — das mag ja noch gehen —, aber als Wertebereich die Menge aller Translationen (die selbst wieder Abbildungen sind). Wir nennen diese Abbildung wieder T – ohne Index. Bezeichnen wir die Menge aller Translationen mit \mathcal{T} , so haben wir $T : \mathbb{R}^n \rightarrow \mathcal{T}$, $\mathbf{a} \mapsto T_{\mathbf{a}}$. Nun machen wir uns klar, dass die Verkettung $T_{\mathbf{a}} \circ T_{\mathbf{b}}$ von zwei Translationen wieder eine Translation, nämlich $T_{\mathbf{c}}$ mit $\mathbf{c} := \mathbf{a} + \mathbf{b}$ ist¹³, kurz

$$T_{\mathbf{a}} \circ T_{\mathbf{b}} = T_{\mathbf{a}+\mathbf{b}} \quad \forall \mathbf{a}, \mathbf{b} \in \mathbb{R}^n.$$

Diese Beziehung gibt einen einfach zu deutenden Sachverhalt wieder: Wenn man zwei Translationen hintereinander ausführt, so hätte man das Resultat auch erhalten, wenn man um nur einen Vektor verschiebt — nämlich um die Summe der beiden Translationsvektoren.

¹³eigentlich mit $\mathbf{c} = \mathbf{b} + \mathbf{a}$ — die Translation $T_{\mathbf{b}}$ wird zuerst ausgeführt. Wegen $\mathbf{b} + \mathbf{a} = \mathbf{a} + \mathbf{b}$ ist unsere Aussage aber richtig.

Abbildung 1.20: $\vec{PQ} + \vec{QR} = \vec{PR}$

Daher ist (\mathcal{T}, \circ) ebenfalls (wie der \mathbb{R}^n) eine abelsche Gruppe, man spricht von der *Translationsgruppe*¹⁴. Insbesondere sind offensichtlich alle Translationen Bijektionen des \mathbb{R}^n . Die Umkehrabbildung von $T_{\mathbf{a}}$ ist $T_{-\mathbf{a}}$. Vielleicht erinnern Sie sich: In Aufgabe I.44 hatten wir erkannt, dass alle Drehungen der Ebene um den Ursprung ebenfalls eine Gruppe mit der Verkettung als binäre Verknüpfung ist — im Gegensatz zu der Menge aller Spiegelungen an Geraden.

Jetzt kommt dem Nullvektor \mathbf{o} auch eine besondere Bedeutung zu: Die zugehörige Translation ist die *Identität*.

Translationen beschreiben die einfachsten *Bewegungen*. Punktmenge(n) (z.B. irgendein geometrisches Gebilde) werden ohne jegliche Formänderung verschoben — Abstände und Winkel bleiben gleich.

Die Addition von Vektoren kann auch gut veranschaulicht werden, wenn man das Konzept eines *Verbindungsvektors* (s. Abb. 1.12) wieder aufgreift. Ist P der Anfangs- und Q der Endpunkt des Pfeils, so wird häufig die Notation $\mathbf{x} = \vec{PQ}$ für den Verbindungsvektor verwendet. In Abb. 1.20 wird diese Pfeilschreibweise benutzt. Sie veranschaulicht dieselbe Aussage wie Abb. 1.19.

1.4 Multiplikation von Vektoren mit Skalaren

Wenn man zu irgendeinem *Skalar*¹⁵ λ eine Abbildung $(x_1, x_2) \mapsto (\lambda x_1, \lambda x_2)$ bzw. $(x_1, x_2, x_3) \mapsto (\lambda x_1, \lambda x_2, \lambda x_3)$ betrachtet, so hat sich bei positivem $\lambda > 0$ die Länge¹⁶ des Vektors \mathbf{x} , nicht aber dessen Richtung verändert. Ist $\lambda < 0$, so kehrt sich die Richtung eines jeden Vektors um. Für $\lambda > 1$ werden Vektoren *gedehnt*, für $0 < \lambda < 1$ *kontrahiert*, s. Abb. 1.21.

¹⁴Dass $T_{\mathbf{a}} \circ T_{\mathbf{b}} = T_{\mathbf{b}} \circ T_{\mathbf{a}} = T_{\mathbf{a}+\mathbf{b}}$ kann man sich durch das Parallelogramm in Abb.1.18 veranschaulicht denken.

¹⁵Elemente von \mathbb{R} nennt man zuweilen Skalare, um sie von den Vektoren zu unterscheiden.

¹⁶Zur Definition der Länge kommen wir noch. Aber es kann nicht schaden, schon jetzt eine klare Vorstellung von der *Länge eines Vektors* \mathbf{x} zu haben: Es ist $|\mathbf{x}| := \sqrt{x_1^2 + x_2^2}$ bzw. im Raum $|\mathbf{x}| := \sqrt{x_1^2 + x_2^2 + x_3^2}$

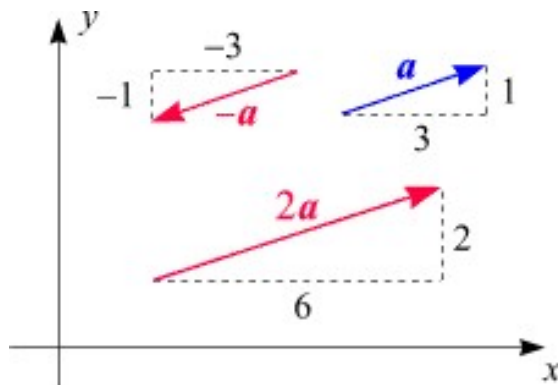


Abbildung 1.21: Skalarmultiplikation

Wie bei Translationen betrachten wir zu jedem Skalar λ die Abbildung $D_\lambda : \mathbb{R}^n \rightarrow \mathbb{R}^n, \mathbf{x} \mapsto \lambda \cdot \mathbf{x}$, wobei wir von einer neuen Schreibweise Gebrauch machen:

Definition 1.3. Ist λ ein Skalar und $\mathbf{x} := (x_1, x_2, \dots, x_n)$ ein Vektor, so wird $\lambda \cdot \mathbf{x}$ durch

$$\lambda \cdot \mathbf{x} := \lambda \cdot (x_1, x_2, \dots, x_n) := (\lambda x_1, \lambda x_2, \dots, \lambda x_n)$$

definiert.

Wir nennen die Abbildung D_λ eine **Streckung**, ganz gleich, wie λ beschaffen ist. D_0 ist dann die *Nullabbildung*, die jedes \mathbf{x} auf den Nullvektor abbildet.

So wie wir bei Translationen die Abbildung $T : \mathbf{a} \mapsto T_{\mathbf{a}}$ betrachtet haben, kann man $D : \lambda \mapsto D_\lambda$ untersuchen. Hier gilt offensichtlich

$$D_{\lambda \cdot \mu} = D_\lambda \circ D_\mu.$$

Die Verkettung von zwei Streckungen um λ und μ ist wieder eine Streckung um $\lambda \cdot \mu$. Sehen Sie die Analogie zu Translationen? So wie (\mathbb{R}^*, \cdot) mit $\mathbb{R}^* = \mathbb{R} \setminus \{0\}$ eine abelsche Gruppe ist, so auch (\mathcal{D}^*, \circ) mit $\mathcal{D} := \{D_\lambda : \lambda \in \mathbb{R}\}$ und $\mathcal{D}^* = \{D_\lambda : \lambda \neq 0\}$.

Bemerkenswert, aber ganz einfach ist die Beobachtung, dass die folgenden Rechenregeln gelten

- **Distributivgesetz 1**

$$(\lambda + \mu) \cdot \mathbf{x} = \lambda \cdot \mathbf{x} + \mu \cdot \mathbf{x}, \quad \forall \mathbf{x} \in \mathbb{R}^n, \lambda, \mu \in \mathbb{R}$$

- **Distributivgesetz 2**

$$\lambda \cdot (\mathbf{x} + \mathbf{y}) = \lambda \cdot \mathbf{x} + \lambda \cdot \mathbf{y} \quad \forall \mathbf{x}, \mathbf{y} \in \mathbb{R}^n, \lambda \in \mathbb{R},$$

- **Assoziativgesetz**

$$(\lambda \cdot \mu) \cdot \mathbf{x} = \lambda \cdot (\mu \cdot \mathbf{x}) \quad \forall \lambda, \mu \in \mathbb{R}, \mathbf{x} \in \mathbb{R}^n,$$

- $1 \cdot \mathbf{x} = \mathbf{x}$.

1.4.1 Vektorraum allgemein

Diese Gesetze werden in einem allgemeinen Vektorraum als eine der grundlegenden Eigenschaften gefordert. Zusammen mit der Eigenschaft „abelsche Gruppe“ von $(\mathbb{R}^n, +)$ machen sie den $(\mathbb{R}^n, +, \cdot)$ zu einem **Vektorraum** über dem Körper \mathbb{R} der reellen Zahlen. Genauer:

Definition 1.4. Sei $(V, +)$ eine abelsche Gruppe mit neutralem Element \mathbf{o} . Ferner sei eine Skalarmultiplikation¹⁷ $(\lambda, \mathbf{v}) \in \mathbb{R} \times V \mapsto \lambda \cdot \mathbf{v}$ erklärt, die mit V an Stelle von \mathbb{R}^n den eben genannten Rechenregeln genügen. Dann heißt $(V, +, \cdot)$ ein **Vektorraum** über dem Körper der reellen Zahlen¹⁸. Die Elemente $\mathbf{v} \in V$ heißen Vektoren, die Elemente λ des Körpers heißen Skalare.

Mit derselben Technik wie in einem Körper kann man übrigens in jedem Vektorraum $0 \cdot \mathbf{x} = \mathbf{o}$ folgern, wobei natürlich \mathbf{o} das neutrale Element bzgl. der Addition ist. Auch $\lambda \cdot \mathbf{o} = \mathbf{o}$ gilt für alle $\lambda \in \mathbb{R}$. Zur Erinnerung: Zum Beweis wende man die Distributivgesetze $(0 + 0) \cdot \mathbf{x} = 0 \cdot \mathbf{x} + 0 \cdot \mathbf{x}$ und $\lambda \cdot (\mathbf{o} + \mathbf{o}) = \lambda \cdot \mathbf{o} + \lambda \cdot \mathbf{o}$ an.

1.4.2 Ein Funktionenraum als Vektorraum

Können Sie sich vorstellen, dass man Funktionen als Elemente eines Vektorraums, also als „Vektoren“ deuten kann? Und entsprechend mit ihnen rechnet?

Hierzu betrachten wir die Menge aller Funktionen von \mathbb{R} nach \mathbb{R} , also alle Ihnen hoffentlich immer vertrauter werdenden reellen Funktionen. Diese Menge nennen wir V — weil wir sie jetzt mit einer Vektorraum-Struktur versehen werden. Wie addieren wir zwei Funktionen $f, g \in V$? Unter $h := f + g$ verstehen wir die Funktion $h : x \mapsto f(x) + g(x)$. Ferner können wir eine „Skalarmultiplikation“ $h := \lambda \cdot f, \lambda \in \mathbb{R}$, durch $h : x \mapsto \lambda f(x)$ definieren. Damit wird V in der Tat zu einem Vektorraum! Sehen Sie die Ähnlichkeit mit dem \mathbb{R}^n , wo „komponentenweise“ addiert und skalar-multipliziert wurde? Eine „Komponente“ einer Funktion f ist gerade $f(x)$ — zu jedem x gibt es eine Komponente! Ja, Sie hätten einen Vektor $\mathbf{v} \in \mathbb{R}^n$ auch als eine Funktion auffassen können, nämlich als eine Funktion $f : \{1, 2, \dots, n\} \rightarrow \mathbb{R}$. Wenn Sie diese Analogie verstanden haben, haben Sie schon ein gutes Gefühl für mathematische Strukturen entwickelt!

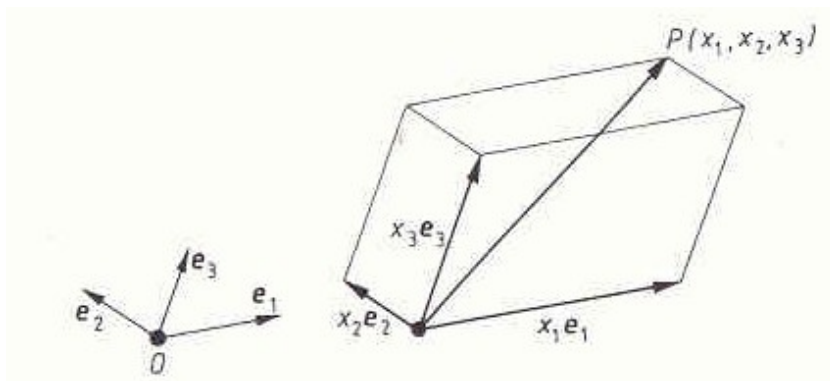
Welche Funktion $f : \mathbb{R} \rightarrow \mathbb{R}$ übernimmt jetzt die Rolle des neutralen Elements \mathbf{o} (bzgl. der Addition)? Richtig! Die „Nullfunktion“, das ist die Funktion, die jedem x den Funktionswert $y = 0$ zuordnet!

1.4.3 Linearkombinationen

Der Begriff *Linearkombination* ist für das Folgende zentral.

¹⁷Nicht zu verwechseln mit dem später eingeführten Skalarprodukt von zwei Vektoren.

¹⁸Es gibt auch Vektorräume über anderen Körpern, z.B. über \mathbb{C} .

Abbildung 1.22: $\mathbf{x} = x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + x_3\mathbf{e}_3$

Definition 1.5. Sind $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_m$ Vektoren (des \mathbb{R}^n) und c_1, c_2, \dots, c_m Skalare (aus \mathbb{R}), so heißt der Vektor

$$\mathbf{v} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \dots + c_m\mathbf{v}_m$$

Linearkombination der \mathbf{v}_j und die c_j heißen **Koeffizienten** der Linearkombination, $j = 1, 2, \dots, m$.

Ich greife noch einmal die in den Abb. 1.7–1.16 skizzierten Koordinatensysteme auf. Die Koordinatenachsen haben eine Richtung. Die Vektoren dieser Richtung mit Länge 1 heißen **Einheitsvektoren** und werden mit $\mathbf{e}_j, j = 1, 2, 3$, bezeichnet. Jetzt kann ein beliebiger Vektor $\mathbf{x} = (x_1, x_2, x_3)$ als *Linearkombination*

$$\mathbf{x} = x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + x_3\mathbf{e}_3$$

mit den Komponenten x_j als Koeffizienten dargestellt werden, s. Abb. 1.22.

Dabei müssen die Koordinatenachsen nicht zwingend senkrecht aufeinander stehen, siehe Abb. 1.23 für ein „schiefes“ (neben einem kartesischen) ebenes Koordinatensystem.

Bei den Begriffen *Basis* und *lineare Unabhängigkeit* (s. Kap. 1.5) spielt die *Eindeutigkeit* einer Linearkombination eine Rolle. Man sagt, dass ein Vektor \mathbf{v} eine eindeutige Linearkombination aus Vektoren $\mathbf{v}_j, j = 1, 2, \dots, m$ besitzt, wenn es genau ein m -Tupel (c_1, c_2, \dots, c_m) von Linearkoeffizienten mit $\mathbf{v} = \sum_{j=1}^m c_j\mathbf{v}_j$ gibt.

Hat man m Nahrungsmittel (NM) und n Nährstoffe (NS), so kann man jedem NM einen „Nährstoffvektor“ $\mathbf{v}_j, j = 1, 2, \dots, m$ zuordnen, dessen k -te Komponente angibt, wieviele Mengeneinheiten (ME) des k -ten NS eine ME des j -ten NM besitzt. Dann kann man einen Einkaufskorb als Linearkombination \mathbf{v} der $\mathbf{v}_j, j = 1, 2, \dots, m$ ansehen, wenn man an den NS-Mengen interessiert ist. Die Linearkoeffizienten geben an, wieviele ME der NM sich im Einkaufskorb befinden.

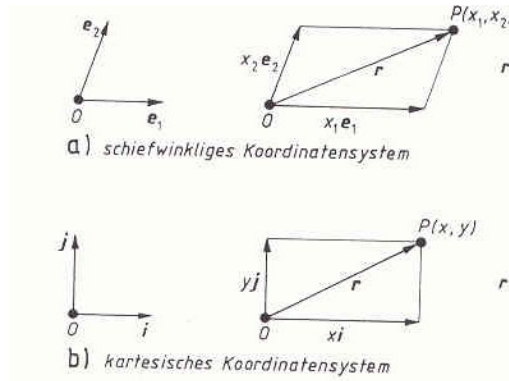


Abbildung 1.23: Schiefes Koordinatensystem

1.4.4 Spezielle Linearkombinationen

Sind \mathbf{a} und \mathbf{b} zwei Vektoren, so ist $\mathbf{c} := \frac{1}{2}\mathbf{a} + \frac{1}{2}\mathbf{b}$ eine spezielle *Linearkombination* von \mathbf{a} und \mathbf{b} , auch (**arithmetischer**) **Mittelwert** von \mathbf{a} und \mathbf{b} genannt. Diesen Mittelwert kann man ganz einfach geometrisch interpretieren: Fasst man \mathbf{a} und \mathbf{b} als Ortsvektoren zweier Punkte A und B auf, so ist \mathbf{c} der zum *Mittelpunkt* der Strecke \overline{AB} gehörende Ortsvektor. Das macht man sich leicht an Hand einer Zeichnung klar!

Dieser geometrische Sachverhalt wird vielleicht deutlicher, wenn wir $\frac{1}{2}\mathbf{a} + \frac{1}{2}\mathbf{b} = \mathbf{a} + \frac{\mathbf{b}-\mathbf{a}}{2}$ und die Interpretation von $\mathbf{b} - \mathbf{a}$ als Verbindungsvektor beachten. Jetzt kann man auch einfach sehen, dass durch

$$\{\mathbf{a} + \lambda(\mathbf{b} - \mathbf{a}) : 0 \leq \lambda \leq 1\}$$

gerade die gesamte Verbindungsstrecke \overline{AB} und durch

$$g_{AB} := \{\mathbf{a} + \lambda(\mathbf{b} - \mathbf{a}) : \lambda \in \mathbb{R}\}$$

die durch A und B verlaufenden Gerade gegeben ist. Wegen $\mathbf{a} + \lambda(\mathbf{b} - \mathbf{a}) = (1 - \lambda)\mathbf{a} + \lambda\mathbf{b}$ ist also auch jeder Punkt der Verbindungsgeraden als spezielle Linearkombination von \mathbf{a} und \mathbf{b} darstellbar.

Sie haben vielleicht gemerkt, dass ich nicht mehr konsequent zwischen Punkt A und Vektor \mathbf{a} bzw. einem Punkt der Geraden g_{AB} und dem zugehörigen Ortsvektor unterschieden habe. Das ist viel bequemer!

Jetzt betrachten wir drei Punkte A, B und C , die wir nun nicht mehr von ihren Ortsvektoren \mathbf{a}, \mathbf{b} und \mathbf{c} unterscheiden wollen. Wie kann man den *Mittelwert*

$$\mathbf{M} = \frac{1}{3}\mathbf{a} + \frac{1}{3}\mathbf{b} + \frac{1}{3}\mathbf{c}$$

bzw.

$$\mathbf{M} = \frac{1}{3}(\mathbf{a} + \mathbf{b} + \mathbf{c})$$

geometrisch interpretieren? Ja: \mathbf{M} ist gerade der Schnittpunkt der drei Seitenhalbierenden des zugehörigen Dreiecks¹⁹. Das sieht man so ein: Die Seitenmittelpunkte sind $M_{ab} := \frac{1}{2}(\mathbf{a} + \mathbf{b})$, $M_{ac} := \frac{1}{2}(\mathbf{a} + \mathbf{c})$, $M_{bc} := \frac{1}{2}(\mathbf{b} + \mathbf{c})$. Wegen $\frac{1}{3}(\mathbf{a} + \mathbf{b} + \mathbf{c}) = \frac{2}{3}M_{ab} + \frac{1}{3}\mathbf{c}$ liegt \mathbf{M} auf der Seitenhalbierenden durch M_{ab} und \mathbf{c} ²⁰. Aus dem selben Grund liegt \mathbf{M} auf jeder Seitenhalbierenden. Wir haben mit Hilfe der Vektorrechnung bewiesen, dass sich die drei Seitenhalbierenden eines Dreiecks in einem Punkt schneiden²¹. Dies ist ein typisches Beispiel für die *Analytische Geometrie*, in der geometrische Probleme mit Hilfe der Vektorrechnung *analysiert* werden.

1.5 Lineare Unabhängigkeit, Basis

1.5.1 Der 3-dimensionale Fall

Ich diskutiere zunächst den dreidimensionalen Fall. Übertragungen auf den ebenen Fall sind offensichtlich, bzw. werden in Kap. 1.5.3 vorgenommen.

In einem kartesischen Koordinatensystem sind die Koordinatenachsen durch die drei Einheitsvektoren $\mathbf{e}_j, j = 1, 2, 3$, (in \mathbb{R}^3) gegeben. Die Schreibweise $\mathbf{a} = (a_1, a_2, a_3)$ für einen Vektor besagt gerade $\mathbf{a} = a_1\mathbf{e}_1 + a_2\mathbf{e}_2 + a_3\mathbf{e}_3$ — \mathbf{a} ist auf eindeutige Weise als Linearkombination der $\mathbf{e}_j, j = 1, 2, 3$ darstellbar mit den Komponenten von \mathbf{a} als Linearkoeffizienten.

Was aber, wenn man die Einheitsvektoren $\mathbf{e}_1 = (1, 0, 0)$, $\mathbf{e}_2 = (0, 1, 0)$ und $\mathbf{e}_3 = (0, 0, 1)$ durch drei beliebige Vektoren $\mathbf{v}_1, \mathbf{v}_2$ und \mathbf{v}_3 ersetzt? Diese lassen sich bzgl. eines kartesischen Koordinatensystems mit Einheitsvektoren $\mathbf{e}_j, j = 1, 2, 3$ darstellen als $\mathbf{v}_1 = (v_{11}, v_{21}, v_{31}), \dots, \mathbf{v}_3 = (v_{13}, v_{23}, v_{33})$. Die drei Vektoren sind also durch eine *rechteckige Tabelle* mit drei Zeilen und drei Spalten mit reellen Einträgen gegeben, wobei \mathbf{v}_j die *j*-te *Spalte* der Tabelle ist²²:

$$\begin{pmatrix} v_{11} & v_{12} & v_{13} \\ v_{21} & v_{22} & v_{23} \\ v_{31} & v_{32} & v_{33} \end{pmatrix} \quad (1.1)$$

Definition 1.6. (Zur allg. Definition einer Basis siehe Def. 1.10.) Drei Vektoren $\mathbf{v}_1, \mathbf{v}_2$ und \mathbf{v}_3 bilden eine **Basis** von \mathbb{R}^3 , wenn sich jeder Vektor $\mathbf{a} \in \mathbb{R}^3$ eindeutig als Linearkombination

$$\mathbf{a} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + c_3\mathbf{v}_3, \quad c_j \in \mathbb{R}, \quad j = 1, 2, 3$$

der $\mathbf{v}_j, j = 1, 2, 3$, darstellen lässt.

Es stellt sich die Frage, ob man der obigen 3×3 -Tabelle ansehen kann, dass es sich bei $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ um eine Basis handelt. Ist $\mathbf{a} = \mathbf{o}$ der Nullvektor, so ist dieser natürlich immer als Linearkombination von drei Vektoren darstellbar: man setze nur $c_j = 0, j = 1, 2, 3$, für alle drei *Linearkoeffizienten*. Man spricht dann von der *trivialen Linearkombination des Nullvektors*. Bilden $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$

¹⁹auch *Schwerpunkt* des Dreiecks!

²⁰Es folgt sogar, dass \mathbf{M} die Seitenhalbierenden im Verhältnis 2:1 teilt. Siehe Übungen.

²¹und auch das Teilverhältnis 2:1, mit dem sich die Seitenhalbierenden schneiden, ermittelt.

²²Das hat mit späteren Matrix-Rechnungen zu tun und deutet schon an, dass das Rechnen mit *Spaltenvektoren* meist bequemer ist.

eine Basis, so muss sich natürlich auch der Nullvektor *eindeutig* als (triviale) Linearkombination der \mathbf{v}_j darstellen lassen. In diesem Fall einer Basis gibt es keine weitere (*nichttriviale*) Linearkombination des Nullvektors.

Beispiel: Jetzt kann man sofort sehen, dass $\mathbf{v}_1 := (1, 0, 0)$, $\mathbf{v}_2 := (0, 1, 0)$, $\mathbf{v}_3 := (1, 1, 0)$ keine Basis des \mathbb{R}^3 bilden, weil durch $\mathbf{o} = \mathbf{v}_1 + \mathbf{v}_2 - \mathbf{v}_3$ eine nichttriviale Linearkombination gegeben ist.

1.5.2 Der n-dimensionale Fall

Wir definieren gleich allgemein für den \mathbb{R}^n :

Definition 1.7. *Irgendwelche k Vektoren $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k$ des \mathbb{R}^n bilden ein **linear unabhängiges System**, falls es keine nichttriviale Linearkombination des Nullvektors durch diese k Vektoren gibt, d.h. falls gilt*

$$\mathbf{o} = \sum_{j=1}^k c_j \mathbf{v}_j \implies c_j = 0, j = 1, 2, \dots, k.$$

*Ein nicht linear unabhängiges System heißt **linear abhängig**.*

Bemerkungen: a) *Lineare Unabhängigkeit* ist *keine* Eigenschaft, die einem einzelnen Vektor \mathbf{v}_j , sondern eine, die der Gesamtheit von Vektoren als Menge (hier „System“ genannt) zukommt.

b) Noch stehen k und n in keiner Beziehung zu einander. Wir werden aber sehen, dass $k \leq n$ sein muss, wenn lineare Unabhängigkeit besteht.

c) Statt „System“ könnte man auch von einer „Menge“ $\{\mathbf{v}_1, \dots, \mathbf{v}_k\}$ reden. Das wird nur problematisch, wenn man z.B. das von $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3 := \mathbf{v}_1$ gebildete System als linear abhängig bezeichnen will, weil $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_1\}$ keine Menge ist.

Ein aus k Vektoren $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k$ des \mathbb{R}^n gebildete System ist also **linear abhängig**, falls es nicht linear unabhängig ist, wenn sich also der Nullvektor *nichttrivial* durch die $\mathbf{v}_j, j = 1, 2, \dots, k$ linear kombinieren lässt, wenn also eine Darstellung

$$\mathbf{o} = \sum_{j=1}^k c_j \mathbf{v}_j$$

mit mindestens einem Linearkoeffizienten $c_j \neq 0$ existiert.

Beispiel: $n = 3, k = 3, \mathbf{v}_1 = (1, 0, 0), \mathbf{v}_2 = (0, 1, 0), \mathbf{v}_3 = (1, 1, 0)$. Diese drei Vektoren bilden ein linear abhängiges System, weil $\mathbf{v}_3 = \mathbf{v}_1 + \mathbf{v}_2$ bzw. $\mathbf{o} = \mathbf{v}_1 + \mathbf{v}_2 - \mathbf{v}_3$ — es gibt eine nichttriviale Linearkombination des Nullvektors mit Linearkoeffizienten $c_1 = 1, c_2 = 1, c_3 = -1$.

Natürlich bilden diese drei Vektoren auch keine Basis des \mathbb{R}^3 , da sich $\mathbf{e}_3 = (0, 0, 1)$ nicht durch die $\mathbf{v}_j, j = 1, 2, 3$ linear kombinieren lässt.

Satz 1.8. *Das System $\{\mathbf{v}_1, \dots, \mathbf{v}_k\}$ ist genau dann linear abhängig, wenn sich irgendeiner der k Vektoren $\mathbf{v}_j, j = 1, 2, \dots, k$ durch die anderen linear kombinieren lässt.*

Beweis: Sei $\mathbf{v}_k = c_1\mathbf{v}_1 + \cdots + c_{k-1}\mathbf{v}_{k-1}$ linear kombinierbar durch die ersten $k - 1$ Vektoren des Systems. Dann gilt

$$\mathbf{o} = c_1\mathbf{v}_1 + \cdots + c_{k-1}\mathbf{v}_{k-1} - \mathbf{v}_k.$$

Dies ist eine nichttriviale Linearkombination des Nullvektors \mathbf{o} , da der letzte Linearkoeffizient $c_k := -1 \neq 0$. Auf den Beweis der Umkehrung verzichte ich hier. ■

Jetzt spielen wir ein wenig mit dem Begriff *linear unabhängiges System*:

$k = 1$: Ein nur aus einem Vektor \mathbf{v}_1 bestehendes System ist linear unabhängig genau dann, wenn $\mathbf{v}_1 \neq \mathbf{o}$. Ist Ihnen dies klar? Beachten Sie bitte, dass $\mathbf{o} = \lambda\mathbf{o}$ für jedes $\lambda \in \mathbb{R}$.

$k = 2$: Ein aus zwei Vektoren $\mathbf{v}_1, \mathbf{v}_2$ bestehendes System ist linear unabhängig genau dann, wenn keiner der beiden Vektoren Vielfaches des anderen ist. Denn wenn z.B. $\mathbf{v}_1 = \lambda\mathbf{v}_2$ mit einem $\lambda \in \mathbb{R}$, so gilt $\mathbf{v}_1 - \lambda\mathbf{v}_2 = \mathbf{o}$, also $c_1\mathbf{v}_1 + c_2\mathbf{v}_2 = \mathbf{o}$ mit $c_1 = 1 (\neq 0)$ und $c_2 = -\lambda$ — wir haben eine nichttriviale Linearkombination von \mathbf{o} und somit ein linear abhängiges System. Nun nehmen wir umgekehrt an, dass \mathbf{v}_1 und \mathbf{v}_2 ein linear abhängiges System ist, d.h. dass $c_1\mathbf{v}_1 + c_2\mathbf{v}_2 = \mathbf{o}$ mit $c_1 \neq 0$ oder $c_2 \neq 0$ gilt. Ich nehme $c_1 \neq 0$ an. Dann können wir die Gleichung $c_1\mathbf{v}_1 + c_2\mathbf{v}_2 = \mathbf{o}$ durch c_1 dividieren und erhalten $\mathbf{v}_1 = \lambda\mathbf{v}_2$ mit $\lambda = -\frac{c_2}{c_1}$, d.h. \mathbf{v}_1 ist ein Vielfaches von \mathbf{v}_2 . Statt korrekt zu sagen: „Das aus \mathbf{v}_1 und \mathbf{v}_2 bestehende System ist linear (un)abhängig“, werden wir meist kürzer (aber sprachlich nicht korrekt) sagen: „Die beiden Vektoren $\mathbf{v}_1, \mathbf{v}_2$ sind linear (un)abhängig“.

Die lineare Unabhängigkeit des aus \mathbf{v}_1 und \mathbf{v}_2 bestehenden Systems bedeutet also nichts anderes als dass sie im Sinne der folgenden Definition nicht *parallel* sind:

Definition 1.9. Zwei Vektoren²³ \mathbf{a} und \mathbf{b} des \mathbb{R}^n heißen genau dann **parallel**, wenn sie ein linear abhängiges System bilden, d.h., falls einer von ihnen Vielfaches des anderen Vektors ist²⁴.

Im \mathbb{R}^n bilden die n **Einheitsvektoren** $\mathbf{e}_j = (0, \dots, 0, 1, 0, \dots, 0)$ (eine 1 an der j -ten Position) ein linear unabhängiges System. Denn es ist $\sum_{j=1}^n c_j\mathbf{e}_j = (c_1, \dots, c_n) = \mathbf{o}$ genau dann, wenn alle $c_j = 0, j = 1, 2, \dots, n$. Nicht so einfach einzusehen ist, dass ein aus $k > n$ Vektoren des \mathbb{R}^n bestehendes System niemals linear unabhängig sein kann.

Die Einheitsvektoren $\mathbf{e}_j, j = 1, 2, \dots, n$ des \mathbb{R}^n bilden eine *Basis* des \mathbb{R}^n in folgendem Sinne (in Verallgemeinerung von Def. 1.6):

Definition 1.10. Ein aus den Vektoren $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n$ des \mathbb{R}^n bestehendes System heißt eine **Basis** des \mathbb{R}^n , falls sich jeder Vektor $\mathbf{x} \in \mathbb{R}^n$ eindeutig durch $\mathbf{x} = \sum_{j=1}^n c_j\mathbf{v}_j$ mit Linearkoeffizienten $c_j, j = 1, 2, \dots, n$ linear kombinieren lässt.

Jetzt formuliere ich ohne Beweis

²³Wir verlangen nicht, dass \mathbf{a} und \mathbf{b} verschieden vom Nullvektor \mathbf{o} sein sollen. Es ergibt sich, dass der Nullvektor parallel zu jedem Vektor ist.

²⁴Wenn keiner der beiden Vektoren der Nullvektor ist, so ist jeder Vektor Vielfaches des anderen.

Satz 1.11. n Vektoren $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n$ bilden eine Basis des \mathbb{R}^n genau dann, wenn sie ein linear unabhängiges System bilden.

Für später halte ich die folgende Fragestellung fest: Welche Eigenschaft muss die aus den Komponenten aller $\mathbf{v}_j = (v_{1j}, v_{2j}, \dots, v_{nj})$ bestehende $n \times n$ -Tabelle (später *Matrix* genannt)

$$\begin{pmatrix} v_{11} & v_{12} & \cdots & v_{1n} \\ v_{21} & v_{22} & \cdots & v_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ v_{n1} & v_{n2} & \cdots & v_{nn} \end{pmatrix}$$

haben, damit wir es mit einer Basis des \mathbb{R}^n zu tun haben?

Bemerkung: Der Begriff *Basis* ist auch für andere Vektorräume als der \mathbb{R}^n wichtig. Die Anzahl der Vektoren, die eine Basis enthält, heißt dann *Dimension* des Vektorraums. In diesem Sinne ist die Dimension von \mathbb{R}^n gerade n , und die Sprechweise „ n -dimensionaler Raum \mathbb{R}^n “ ist gerechtfertigt.

1.5.3 Der ebene Fall

Für $n = 2$ besagt der Satz 1.11, dass sich jeder Vektor $\mathbf{a} = (a_1, a_2)$ auf eindeutige Weise als Linearkombination $\mathbf{a} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2$ darstellen lässt, wenn das aus $\mathbf{v}_1 = (v_{11}, v_{21})$ und $\mathbf{v}_2 = (v_{12}, v_{22})$ bestehende System linear unabhängig ist, wenn also \mathbf{v}_1 und \mathbf{v}_2 nicht parallel sind.

Ein lineares Gleichungssystem

Es ist nicht sehr schwierig, den Satz 1.11 für den ebenen Fall $n = 2$ zu beweisen. Hierzu muss man sich klar machen, dass die eindeutige Linearkombination $\mathbf{a} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2$ mit der eindeutigen Lösbarkeit eines linearen Gleichungssystems gleichwertig ist. Man sucht ja zu gegebenem \mathbf{a} (und $\mathbf{v}_1, \mathbf{v}_2$) die beiden Linearkoeffizienten c_1 und c_2 . Wie häufig in der Mathematik verwenden wir den Buchstaben x für das Unbekannte, also hier x_1 an Stelle von c_1 und x_2 an Stelle von c_2 . Dann haben wir es mit zwei linearen Gleichungen

$$v_{11}x_1 + v_{12}x_2 = a_1, \quad v_{21}x_1 + v_{22}x_2 = a_2 \tag{1.2}$$

zu tun — einem **linearen Gleichungssystem**. Wir nennen die Tabelle

$$V := \begin{pmatrix} v_{11} & v_{12} \\ v_{21} & v_{22} \end{pmatrix}$$

die **Koeffizientenmatrix** von (1.2) und den Ausdruck

$$\Delta(V) := v_{11}v_{22} - v_{12}v_{21} \tag{1.3}$$

die **Determinante**²⁵. Nun kann man mit elementarer Rechnung zeigen

Satz 1.12. *Die folgenden Aussagen sind gleichwertig:*

- (L1) *Das System der beiden Vektoren \mathbf{v}_1 und \mathbf{v}_2 ist linear unabhängig.*
- (L2) *Die Determinante $\Delta(V)$ ist ungleich Null.*
- (L3) *Das lineare Gleichungssystem (1.2) hat eine eindeutige Lösung.*

Mehr hierzu in den Übungen.

Hier wollen wir diesen Sachverhalt in einem Spezialfall $\mathbf{v}_1 = (1, 1)$ untersuchen. (L1) besagt gerade, dass $\mathbf{v}_2 = (v_{12}, v_{22})$ kein Vielfaches von \mathbf{v}_1 ist, d.h. (L1) ist mit $v_{12} \neq v_{22}$ äquivalent.

Es ist $\Delta(V) = v_{11}v_{22} - v_{12}v_{21} = v_{22} - v_{12}$. Also sind (L1) und (L2) äquivalent.

Das lineare Gleichungssystem (1.2) lautet

$$x_1 + v_{12}x_2 = a_1, \quad x_1 + v_{22}x_2 = a_2.$$

Subtraktion der Gleichungen (die erste von der zweiten) ergibt (beachte, dass hier $\Delta(V) = v_{22} - v_{12}$)

$$\Delta(V)x_2 = a_2 - a_1.$$

Diese Gleichung besitzt nur für $\Delta(V) \neq 0$ eine eindeutige Lösung x_2 : Ist $\Delta(V) = 0$, so gibt es keine Lösung, wenn $a_1 \neq a_2$, und unendlich viele Lösungen, wenn $a_1 = a_2$ ist.

Die Vektoren \mathbf{v}_1 und \mathbf{v}_2 sind die Spalten der „Matrix“ V . Um auszudrücken, dass die Determinante von diesen beiden Vektoren abhängt, schreiben wir auch $\Delta(\mathbf{v}_1, \mathbf{v}_2)$ an Stelle von $\Delta(V)$. Wir notieren: Sei $\mathbf{a} = (a_1, a_2)$ und $\mathbf{b} = (b_1, b_2)$. Dann gilt

$$\Delta(\mathbf{a}, \mathbf{b}) = a_1b_2 - a_2b_1. \tag{1.4}$$

Determinante und Flächeninhalt eines Parallelogramms

Aus der Schule dürften Sie noch die Formel für den Flächeninhalt eines Parallelogramms kennen: $F = a \cdot h$, wenn a die Länge einer Seite und h die Länge der Höhe über dieser Seite ist.

Vielleicht überrascht Sie der folgende Sachverhalt:

Satz 1.13. *Seien \mathbf{a} und \mathbf{b} zwei Vektoren und P das Parallelogramm, das von \mathbf{a} und \mathbf{b} „erzeugt“ wird, d.h., \mathbf{a} und \mathbf{b} seien die Kantenvektoren von P . Dann ist $|\Delta(\mathbf{a}, \mathbf{b})| = |a_1b_2 - a_2b_1|$ der Flächeninhalt von P .*

Diesen Satz macht man sich in dem Spezialfall $\mathbf{a} = (a, 0)$ und $\mathbf{b} = (b, h)$ sofort klar: Es ist $\Delta(\mathbf{a}, \mathbf{b}) = a \cdot h$! Dass dieser Sachverhalt auch bei beliebiger Lage des Parallelogramms (oder des Koordinatensystems!) gilt, liegt an einer etwas tiefer liegenden Eigenschaft der Determinante, nämlich der Invarianz gegenüber Koordinatensystemen!

Es gibt aber auch einen einfachen direkten Beweis, der mit Schulmathematik (Strahlensatz u.a.) auskommt. Später, in Satz 2.5 in Kap. 2.2.3 wird dieser Satz mit Hilfe von *Skalarprodukten* gezeigt.

²⁵einer 2×2 -Matrix. Später werden Sie *Determinanten* von $n \times n$ -Matrizen kennen lernen.

1.6 Länge von Vektoren

Wie schon durch eine Fußnote bemerkt, ist die (*euklidische*) Länge eines Vektors \mathbf{x} durch $|\mathbf{x}| := \sqrt{x_1^2 + x_2^2}$ bzw. im Raum durch $|\mathbf{x}| := \sqrt{x_1^2 + x_2^2 + x_3^2}$ gegeben (Anwendung des Satzes von Pythagoras). In der komplexen Zahlenebene entspricht dies dem Betrag $|z|$ einer komplexen Zahl.

Die „Länge“ eines Vektors hängt mit dem Begriff „Winkel“ zusammen, den zwei Vektoren miteinander bilden, weil auf den Satz von Pythagoras zurückgegriffen wird, der ja bekanntlich für ein *rechtwinkliges Dreieck* formuliert wird. Insofern gehört dieser Abschnitt eigentlich in das nächste Kapitel.

Eine Übertragung der Längendefinition auf den \mathbb{R}^n ist offensichtlich:

$$|\mathbf{x}| := \sqrt{x_1^2 + x_2^2 + \cdots + x_n^2},$$

oder kürzer

$$|\mathbf{x}| := \sqrt{\sum_{j=1}^n x_j^2}.$$

Wir notieren das offensichtlich gültige Gesetz

$$|\lambda \cdot \mathbf{x}| = |\lambda| |\mathbf{x}| \quad \forall \lambda \in \mathbb{R}, \mathbf{x} \in \mathbb{R}^n.$$

1.6.1 Abstand zwischen zwei Punkten

Sind P und Q zwei Punkte, so ist ihr Abstand die Länge des Verbindungsvektors. Oder: Wenn wir von zwei (Orts-) Vektoren \mathbf{a} und \mathbf{b} ausgehen, so ist ihr Abstand die euklidische Länge von $\mathbf{a} - \mathbf{b}$. Wir geben gleich die Formel im \mathbb{R}^n an: Sei $\mathbf{a} = (a_1, \dots, a_n)$ und $\mathbf{b} = (b_1, \dots, b_n)$, so ist

$$|\mathbf{a} - \mathbf{b}| = \sqrt{\sum_{j=1}^n (b_j - a_j)^2}$$

der Abstand zwischen \mathbf{a} und \mathbf{b} .

1.6.2 Normierung auf Länge Eins

Häufig kommt es nicht auf die Länge von Vektoren an, sondern nur auf ihre Richtung. Sei $\mathbf{a} \neq \mathbf{o}$. Dann gilt für

$$\hat{\mathbf{a}} := \frac{\mathbf{a}}{|\mathbf{a}|},$$

dass $\hat{\mathbf{a}}$ die Länge Eins hat ($|\hat{\mathbf{a}}| = 1$) und dieselbe Richtung wie \mathbf{a} hat. Man sagt, dass $\hat{\mathbf{a}}$ aus \mathbf{a} durch **Normierung** entstehen.

Jetzt kann man die Parallelität von zwei Vektoren $\mathbf{a} \neq \mathbf{o}$ und $\mathbf{b} \neq \mathbf{o}$ dadurch ausdrücken, dass $\hat{\mathbf{a}} = \hat{\mathbf{b}}$ oder $\hat{\mathbf{a}} = -\hat{\mathbf{b}}$ gilt, wobei $\hat{\mathbf{a}}$ aus \mathbf{a} und $\hat{\mathbf{b}}$ aus \mathbf{b} durch Normierung entsteht.

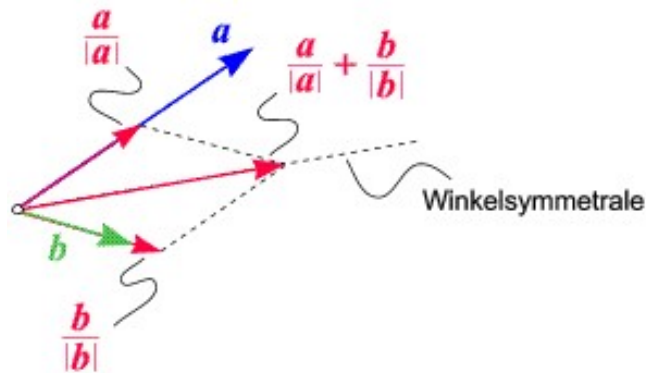


Abbildung 1.24: Winkelsymmetrie

Konstruktion von Winkelhalbierenden

Obwohl wir erst im nächsten Kapitel mehr über Winkel erfahren, möchten wir ein kleines Resultat schon jetzt ernten.

Wir setzen bei der Beobachtung an, dass die Diagonale einer Raute Winkelhalbierende der beiden zugehörigen Winkel ist. Eine Raute ist ein spezielles Parallelogramm (alle Seiten sind gleich lang). Sind \mathbf{a} und \mathbf{b} die „Seitenvektoren“ des Parallelogramms, so ist ja gerade $\mathbf{a} + \mathbf{b}$ der „Diagonalenvektor“. Wenn $|\mathbf{a}| = |\mathbf{b}|$, so ist das Parallelogramm eine Raute und $\mathbf{a} + \mathbf{b}$ ergibt also die Winkelhalbierende.

Nun seien \mathbf{a} und \mathbf{b} zwei beliebige Vektoren, zu denen die Winkelhalbierende (des von ihnen eingeschlossenen Winkels) gesucht ist. Wir machen aus ihnen eine Raute, indem wir \mathbf{a} und \mathbf{b} auf gleiche Länge trimmen, etwa auf Länge Eins. Also definiert

$$\frac{\mathbf{a}}{|\mathbf{a}|} + \frac{\mathbf{b}}{|\mathbf{b}|}$$

die Richtung der Winkelhalbierenden, s. Abb. 1.24.

Beispiel: $\mathbf{a} = (4, 3)$, $\mathbf{b} = (1, 2)$. Dann ist $|\mathbf{a}| = 5$, $|\mathbf{b}| = \sqrt{5}$ und

$$\frac{\mathbf{a}}{|\mathbf{a}|} + \frac{\mathbf{b}}{|\mathbf{b}|} = \left(\frac{4}{5}, \frac{3}{5}\right) + \left(\frac{1}{\sqrt{5}}, \frac{2}{\sqrt{5}}\right).$$

Der Vektor $\left(\frac{4}{5} + \frac{1}{\sqrt{5}}, \frac{3}{5} + \frac{2}{\sqrt{5}}\right)$ ist somit Richtungsvektor der Winkelhalbierenden.

1.7 Geraden

In diesem (Geraden) und dem nächsten Abschnitt (Ebenen) geht es um Anwendungen der Vektorrechnung auf die euklidische Geometrie.

Aus der Schule kennen Sie Geraden wahrscheinlich als Bilder (Graphen) einer Funktion $y = mx + b$, wobei m die Steigung der Geraden und b der Achsenabschnitt auf der y -Achse ist. Genauer ist $\mathcal{G} := \{(x, mx + b) : x \in \mathbb{R}\}$ diese Gerade — als *Punktmenge* des \mathbb{R}^2 . Mit dieser Sichtweise werden Geraden, die parallel zur y -Achse sind, nicht erfasst²⁶. Solche „vertikalen“ Geraden sind durch „ $x = a$ “ oder genauer durch $\mathcal{G} = \{(a, y) : y \in \mathbb{R}\}$ gegeben. Nicht erfasst werden zudem auf diese Weise Geraden im Raum. Dies ist bei dem folgenden Zugang anders:

Wenn zwei verschiedene Punkte A, B einer Geraden \mathcal{G} gegeben sind, kann man deren Verbindungsvektor $\mathbf{v} := \overrightarrow{AB}$ betrachten. Offensichtlich ist die Richtung dieses Vektors (bis auf die Orientierung) für je zwei Punkte der Geraden \mathcal{G} stets dieselbe. Normiert man ihn auf Länge Eins, ist dieser *Richtungsvektor* bis auf das Vorzeichen eindeutig bestimmt, also unabhängig von den betrachteten beiden Punkten. *Parallele* Geraden haben offensichtlich *parallele* Richtungsvektoren (im Sinne von Def. 1.9).

Eine Gerade \mathcal{G} kann nun durch einen Punkt P und einen Richtungsvektor \mathbf{v} (nicht notwendig der Länge Eins) eindeutig bestimmt werden. Jeder Punkt X auf \mathcal{G} kann nämlich durch eine „Koordinate“ λ mittels $X = P + \lambda\mathbf{v}$ beschrieben werden, es ist danach

$$\mathcal{G} = \{P + \lambda\mathbf{v} : \lambda \in \mathbb{R}\}. \quad (1.5)$$

Zu jedem Punkt $X \in \mathcal{G}$ gibt es genau einen „Parameter“ $\lambda \in \mathbb{R}$ mit $X = P + \lambda\mathbf{v}$, ausgeschrieben für $n = 3$ und $X = (x_1, x_2, x_3)$:

$$x_1 = P_1 + \lambda v_1, \quad x_2 = p_2 + \lambda v_2, \quad x_3 = p_3 + \lambda v_3. \quad (1.6)$$

Man nennt (1.5) mit der Kurzform (1.6) die **Parameterdarstellung** oder **Parameterform** von Geraden durch einen reellen Parameter λ . Sie gilt in beliebigen Vektorräumen, nicht nur im \mathbb{R}^n , insbesondere nicht nur in der Ebene. Man stelle sich einen Film vor, in dem bei $\lambda = 0$ beginnend mit einem kontinuierlich wachsenden λ die Punkte $X = P + \lambda\mathbf{v}$ gezeichnet werden. Dann entsteht nach und nach eine Halbgerade mit Anfangspunkt P . Wenn man das gleiche mit kontinuierlich fallendem λ macht, entsteht der fehlende Teil der Geraden. λ ist sozusagen eine Koordinate auf \mathcal{G} , die als Koordinatenachse in P ihren Ursprung hat und dessen Skala durch die Länge von \mathbf{v} gegeben ist.

Während (1.5) die Gerade als Punktmenge beschreibt, kann man die etwas weniger präzise Darstellung (1.6) als die „Parametrisierung“ bezeichnen.

Beispiel: Gegeben seien die beiden Punkte $P = (2, -3), Q = (3, -2)$ in \mathbb{R}^2 . Gesucht ist die Parameterform der Geraden \mathcal{G} durch P und Q . Der Richtungsvektor ist $\mathbf{v} = \overrightarrow{PQ} = (1, 1)$. Setzt man $X = P + \lambda\mathbf{v}$ für einen Punkt auf \mathcal{G} mit $X = (x_1, x_2)$, so ergibt sich die Parametrisierung

$$x_1 = 2 + \lambda, \quad x_2 = -3 + \lambda.$$

(1.5) hat jetzt die Form

$$\mathcal{G} = \{(2 + \lambda, -3 + \lambda) : \lambda \in \mathbb{R}\}.$$

²⁶Die Steigung der Geraden wäre quasi ∞ .

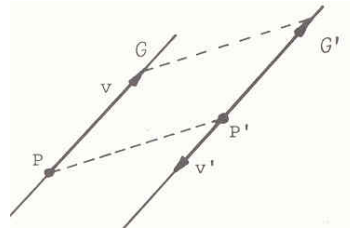


Abbildung 1.25: Parallele Geraden

Ist eine zweite Gerade $\mathcal{G}' = \{P' + \lambda' \mathbf{v}' : \lambda' \in \mathbb{R}\}$ gegeben, so sind diese beiden Geraden genau dann parallel zueinander, wenn die Vektoren \mathbf{v} und \mathbf{v}' linear abhängig sind, s. Abb. 1.25.

Beachten Sie den Zusammenhang mit der Beschreibung einer Verbindungsstrecke und der durch A und B verlaufenden Gerade

$$g_{AB} := \{\mathbf{a} + \lambda(\mathbf{b} - \mathbf{a}) : \lambda \in \mathbb{R}\}$$

aus Kap. 1.4.4.

Beispiele: a) Eine zur y -Achse parallele Gerade hat den zweiten Einheitsvektor $\mathbf{e}_2 = (0, 1)$ als Richtungsvektor. Damit ist die durch $P := (3, 0)$ gehende, zur y -Achse parallele Gerade durch $\mathcal{G} = \{(3, \lambda) : \lambda \in \mathbb{R}\}$ (kurz durch die Parametrisierung $x_1 = 3, x_2 = \lambda$) gegeben.

b) Die z -Achse in \mathbb{R}^3 ist durch $P = (0, 0, 0)$ und Richtungsvektor $\mathbf{v} = (0, 0, 1) = \mathbf{e}_3$ gegeben! Ihre Parameterform ist $\mathcal{G} = \{(0, 0, \lambda) : \lambda \in \mathbb{R}\}$, kurz $x_1 = 0, x_2 = 0, x_3 = \lambda$.

In der Ebene gibt es einen einfachen Zusammenhang zwischen Steigung m und Richtungsvektor $\mathbf{v} = (v_1, v_2)$ einer Geraden: Es ist $m = \frac{v_2}{v_1}$, was nur einen Sinn macht, wenn der Nenner $v_1 \neq 0$. Wenn $v_1 = 0$, so ist der Richtungsvektor parallel zur y -Achse.

1.7.1 Koordinatenform von Geraden in der euklidischen Ebene

Aus der Schule lernten Sie, Geraden durch „(Funktions-)Gleichungen“ zu beschreiben, also durch $y = mx + b$ oder $x = a$. Etwas allgemeiner kann man zu gegebenen reellen Koeffizienten $a, b, c \in \mathbb{R}$ die Punktmenge

$$G := \{(x, y) \in \mathbb{R}^2 : ax + by = c\} \quad (1.7)$$

definieren. Ist G eine Gerade? Nicht notwendigerweise: Wenn $a = 0, b = 0$, aber $c \neq 0$, gilt $G = \emptyset$. Wenn $a = 0, b = 0, c = 0$, also alle Koeffizienten verschwinden, so gilt $G = \mathbb{R}^2$. In allen anderen Fällen ($a \neq 0$ oder $b \neq 0$) ist G in der Tat eine Gerade. Wenn wir das Skalarprodukt von Vektoren kennengelernt haben (Kap. 2), werden wir dem Vektor (a, b) in der Geradengleichung eine geometrische Bedeutung geben können: er steht senkrecht auf der Geraden!

Beispiel: Gegeben seien die beiden Punkte $P := (2, -3), Q := (3, -2)$ in \mathbb{R}^2 . Gesucht ist die Koordinatenform der Geraden \mathcal{G} durch P und Q : Wir hatten dieses Beispiel schon im vorangehenden Kap. 1.7 behandelt und die Parametrisierung

$$x_1 = 2 + \lambda, x_2 = -3 + \lambda, \lambda \in \mathbb{R}$$

berechnet. Jetzt braucht man nur $\lambda = x_1 - 2$ aus der ersten Gleichung zu eliminieren und erhält mit der zweiten Gleichung $x_2 = -3 + \lambda = -3 + x_1 - 2 = x_1 - 5$ oder

$$x_1 - x_2 = 5.$$

Also gilt $\mathcal{G} = \{(x_1, x_2) : ax_1 + bx_2 = c\}$ mit $a = 1, b = -1$ und $c = 5$. Dieses Beispiel lässt sich unschwer verallgemeinern.

1.7.2 Kollinearität von Punkten

Wann sind drei verschiedene Punkte $P, Q, R \in \mathbb{R}^n$ *kollinear*, d.h. wann liegen alle drei auf einer Geraden \mathcal{G} ?

Wir setzen $\mathbf{v} := \vec{PQ}$ und \mathcal{G} als die Gerade durch P mit Richtungsvektor $\mathbf{v} \neq \mathbf{o}$. Nun ist $R \in \mathcal{G}$ genau dann, wenn es ein $\lambda \in \mathbb{R}$ gibt mit $R = P + \lambda\mathbf{v}$. Dann ist aber $\vec{PR} = \lambda\mathbf{v}$, d.h. \vec{PQ} und \vec{PR} sind parallel bzw. — was das gleiche ist — sie sind linear abhängig. Wir fassen zusammen:

Satz 1.14 (Kollinearitätskriterium). *Drei Punkte $P, Q, R \in \mathbb{R}^n$ sind genau dann kollinear, wenn die beiden Richtungsvektoren \vec{PQ} und \vec{PR} linear abhängig sind²⁷.*

1.7.3 Schnitt zweier Geraden

In der Ebene \mathbb{R}^2 erwarten wir, dass sich zwei nicht-parallele Geraden in genau einem Punkt schneiden. In \mathbb{R}^3 (oder erst recht in \mathbb{R}^n mit $n > 3$) muss das nicht sein²⁸.

Wie berechnet man in \mathbb{R}^2 den Schnittpunkt der beiden Geraden

$$\mathcal{G} = \{P + \lambda\mathbf{v} : \lambda \in \mathbb{R}\}$$

und

$$\mathcal{G}' = \{P' + \lambda'\mathbf{v}' : \lambda' \in \mathbb{R}\}?$$

Dies ist einfach: Seien die Koordinaten der beteiligten Größen gegeben, also $P = (p_1, p_2), P' = (p'_1, p'_2), \mathbf{v} = (v_1, v_2)$ und $\mathbf{v}' = (v'_1, v'_2)$. Gesucht ist der Schnittpunkt $S = (s_1, s_2) \in \mathcal{G} \cap \mathcal{G}'$. Wenn $S \in \mathcal{G}$, so muss es ein $x_1 := \lambda$ und ein $x_2 := \lambda'$ geben mit

$$S = P + x_1\mathbf{v} = P' + x_2\mathbf{v}'.$$

²⁷Die „Bevorzugung“ des Punktes P in dieser Definition ist nur scheinbar gegeben.

²⁸Es handelt sich dann um „windschiefe“ Geraden.

Dies führt auf die beiden linearen Gleichungen in den beiden Unbekannten x_1 und x_2

$$p_1 + x_1 v_1 = p'_1 + x_2 v'_1, \quad p_2 + x_1 v_2 = p'_2 + x_2 v'_2.$$

Dieses lineare Gleichungssystem bringen wir in die Form von (1.2):

$$x_1 v_1 - x_2 v'_1 = p'_1 - p_1, \quad x_1 v_2 - x_2 v'_2 = p'_2 - p_2.$$

Die Koeffizientenmatrix hat die Vektoren \mathbf{v} und $-\mathbf{v}'$ in den Spalten. Diese beiden sind linear unabhängig (nach Voraussetzung sind beide Vektoren nicht parallel), ihre Determinante ist ungleich Null und man erhält eine eindeutige Lösung.

1.7.4 Wie löst man ein (ebenes) lineares Gleichungssystem?

Es ist an der Zeit, ein bis zwei Worte zu verlieren, wie man ein lineares Gleichungssystem (1.2), also

$$v_{11}x_1 + v_{12}x_2 = a_1, \quad v_{21}x_1 + v_{22}x_2 = a_2 \tag{1.8}$$

löst. Ich vermute, dass Sie dies wissen. Wirklich interessant ist es erst, einen Algorithmus für beliebige lineare Gleichungssysteme (m Gleichungen mit n Unbekannten; Meist ist $m = n$) zu finden. Stichwort ist der *Gauß'sche Algorithmus*. Hier beruht dieser auf Folgendem: Wenn die Determinante der Koeffizientenmatrix ungleich Null ist, muss mindestens eine der beiden Zahlen, v_{11} oder v_{21} ungleich Null sein. Daher kann man in jedem Fall die erste oder die zweite Gleichung nach x_1 auflösen. Zur Not könnte man die beiden Gleichungen vertauschen, also nehme ich $v_{11} \neq 0$ an. Dann ist

$$x_1 = \frac{a_1 - v_{12}x_2}{v_{11}}.$$

Dies setzt man in die zweite Gleichung ein und erhält nur noch eine einzige lineare Gleichung für x_2 , nämlich

$$\frac{v_{21}}{v_{11}}(a_1 - v_{12}x_2) + v_{22}x_2 = a_2.$$

Der Koeffizient vor x_2 ist $v_{22} - \frac{v_{21}}{v_{11}}v_{12}$ und damit nicht Null, wenn die Determinante nicht verschwindet! Daher kann man nach x_2 auflösen und erhält danach durch Einsetzen von x_2 in die vorherige Gleichung auch x_1 .

Bequemer ist es zuweilen, eine der beiden Gleichungen mit etwas zu multiplizieren und von der andern abzuziehen oder zur anderen zu addieren, um auf diese Weise eine Unbekannte zu „eliminieren“.

1.8 Ebenen im dreidimensionalen Raum

Sind P, Q, R drei nicht kollineare Punkte des \mathbb{R}^n (wichtigster Fall: $n = 3$), so sind die Richtungsvektoren $\mathbf{v} := \vec{PQ}$ und $\mathbf{w} := \vec{PR}$ linear unabhängig. Anschaulich erwarten wir, dass es

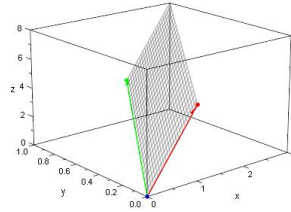


Abbildung 1.26: Die beiden Geradenscharen einer Ebene

genau eine Ebene \mathcal{E} durch diese drei Punkte gibt. Wir werden \mathcal{E} wie schon Geraden durch eine *Parameterdarstellung* beschreiben — dieses Mal sind aber zwei Parameter, die ich mit den griechischen Buchstaben λ und μ bezeichne, notwendig:

$$\mathcal{E} = \{P + \lambda \mathbf{v} + \mu \mathbf{w} : \lambda, \mu \in \mathbb{R}\} \quad (1.9)$$

\mathbf{v} und \mathbf{w} bilden ein sogenanntes **Richtungspaar** der Ebene, auch **Spannvektoren** genannt. Dieses Richtungspaar ist allerdings durch die Ebene keineswegs eindeutig festgelegt!

Wie können wir uns (1.9) klar machen? Ist $\mu = 0$, so erhalten wir die durch P gehende Gerade $\{P + \lambda \mathbf{v} : \lambda \in \mathbb{R}\}$ mit Richtungsvektor \mathbf{v} . Dass diese in \mathcal{E} liegt, verwundert nicht. Ist $\lambda = 0$, so erhalten wir die durch P gehende Gerade $\{P + \mu \mathbf{w} : \mu \in \mathbb{R}\}$ mit Richtungsvektor \mathbf{w} . Auch diese Gerade sollte in \mathcal{E} liegen. Ist μ beliebig, aber fest, so erhalten wir die durch $P + \mu \mathbf{w}$ gehende Gerade $\{(P + \mu \mathbf{w}) + \lambda \mathbf{v} : \lambda \in \mathbb{R}\}$ mit Richtungsvektor \mathbf{v} , also eine ganze Schar von parallelen Geraden, die sämtlich in \mathcal{E} liegen. Genauso gibt es zu jedem festen λ eine durch $P + \lambda \mathbf{v}$ gehende Gerade $\{(P + \lambda \mathbf{v}) + \mu \mathbf{w} : \mu \in \mathbb{R}\}$, die zu $\{P + \mu \mathbf{w} : \mu \in \mathbb{R}\}$ parallel ist. Können Sie sich \mathcal{E} als von zwei Scharen paralleler Geraden „aufgespannter“ Ebene vorstellen? Siehe Abb. 1.26 mit $\mathbf{v} := (1, 0, 5)$ (rot) und $\mathbf{w} := (2, 1, 3)$ (grün).

Man kann sich auch ein Koordinatensystem in der Ebene \mathcal{E} vorstellen: Der Punkt P ist der Ursprung, in dem sich zwei Koordinatenachsen in Richtung \mathbf{v} und \mathbf{w} schneiden, deren Skala durch die Länge von \mathbf{v} und \mathbf{w} gegeben ist. Dann sind λ und μ die Koordinaten eines Punktes der Ebene bzgl. dieses Koordinatensystems. Vergleichen Sie dies mit der Parameterform einer Geraden, wo die Gerade als Koordinatenachse aufgefasst wird, auf der ein Ursprung und eine Skala festgelegt wird.

Es gibt aber noch weitere Geraden in \mathcal{E} , die nicht zu einer der beiden Geradenscharen gehören. Man braucht nur irgendwelche Linearkombination von \mathbf{v} und \mathbf{w} zu bilden und eine Geradenschar mit dieser Linearkombination als Richtungsvektor zu betrachten. Später werden wir sehen, dass es nur auf das *Erzeugnis* $V := \langle \mathbf{v}, \mathbf{w} \rangle$ (s. Def. 1.15) ankommt, d.h., dass man \mathbf{v} und \mathbf{w} durch beliebige zwei linear unabhängige Vektoren aus diesem Erzeugnis V ersetzen kann.

Für $n = 2$ ergibt sich die einzige Ebene $\mathcal{E} = \mathbb{R}^2$, die beiden Vektoren \mathbf{v} und \mathbf{w} bilden eine Basis des \mathbb{R}^2 . Ist $Q = P + \lambda \mathbf{v} + \mu \mathbf{w} \in \mathcal{E}$, so kann man λ und μ als Koordinaten von Q auffassen — bzgl. eines schiefen Koordinatensystems mit Ursprung P und Koordinatenachsen in Richtung \mathbf{v} und \mathbf{w} .

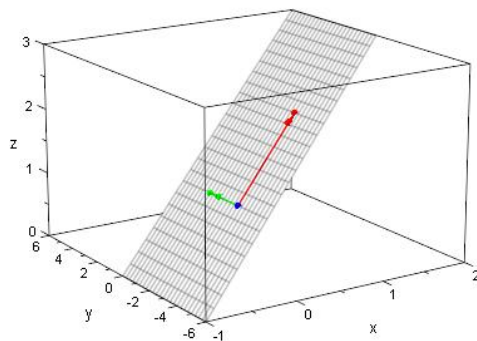


Abbildung 1.27: Ebene

Interessant ist die Ebenenkonstruktion erst für $n \geq 3$.

Beispiel: Gegeben seien die drei Punkte $P = (0, -2, 1)$, $Q = (1, 0, 2)$, $R = (0, 0, 1)$. Mögliche Richtungsvektoren sind $\mathbf{v} = \vec{PQ} = (1, 2, 1)$ und $\mathbf{w} = \vec{PR} = (0, 2, 0)$. Diese beiden Vektoren sind linear unabhängig, da keiner Vielfaches des anderen ist. Nennen wir den allgemeinen Punkt der Ebene $X = (x_1, x_2, x_3)$, so wird dieser durch $X = P + \lambda\mathbf{v} + \mu\mathbf{w}$ mit $\lambda, \mu \in \mathbb{R}$, also durch

$$x_1 = \lambda, \quad x_2 = -2 + 2\lambda + 2\mu, \quad x_3 = 1 + \lambda \quad (1.10)$$

parametrisiert. Durch

$$\mathcal{E} = \{(\lambda, -2 + 2\lambda, 2\mu) : \lambda, \mu \in \mathbb{R}\}$$

ist die Parameterform der Ebene gegeben, siehe Abb. 1.27, die mit MuPAD erstellt wurde. Der rote Pfeil ist der Vektor $\mathbf{v} = \vec{PQ}$, der grüne der Vektor $\mathbf{w} = \vec{PR} = (0, 2, 0)$.

1.8.1 Parallele Ebenen

Ersetzt man in (1.9) den Punkt P durch einen Punkt $P' \notin \mathcal{E}$, lässt das Richtungspaar aber unverändert, so erhält man durch

$$\mathcal{E}' = \{P' + \lambda\mathbf{v} + \mu\mathbf{w} : \lambda, \mu \in \mathbb{R}\}$$

eine zu \mathcal{E} *parallele Ebene*, siehe Abb. 1.28.

Sind umgekehrt

$$\mathcal{E} = \{P + \lambda\mathbf{v} + \mu\mathbf{w} : \lambda, \mu \in \mathbb{R}\}$$

und

$$\mathcal{E}' = \{P' + \lambda\mathbf{v}' + \mu\mathbf{w}' : \lambda, \mu \in \mathbb{R}\}$$

zwei parallele Ebenen, so geht \mathcal{E}' durch die Translation $T_{\mathbf{a}}$ mit $\mathbf{a} = \vec{PP'}$ aus \mathcal{E} hervor, wobei es durchaus sein kann, dass die Richtungspaare verschieden sind, s. Abb. 1.29.

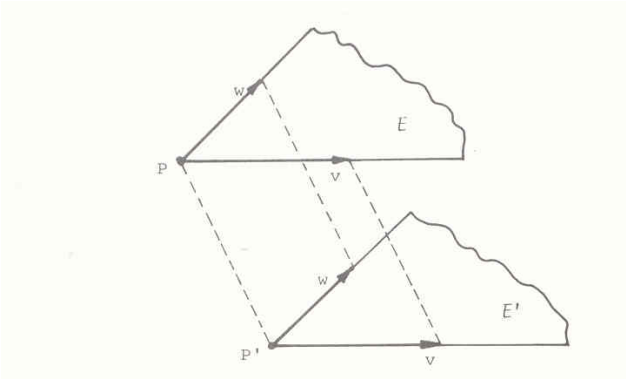


Abbildung 1.28: Die parallelen Ebenen \mathcal{E} und \mathcal{E}'

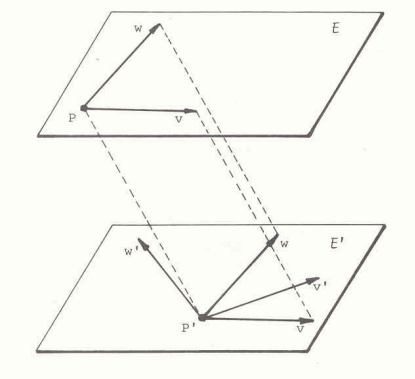


Abbildung 1.29: Die parallelen Ebenen \mathcal{E} und \mathcal{E}'

Es ist $T_a(P+\mathbf{v}) = P'+\mathbf{v} \in \mathcal{E}'$. Daher muss sich \mathbf{v} als Linearkombination von \mathbf{v}' und \mathbf{w}' darstellen lassen. Genauso \mathbf{w} . Alle Linearkombinationen von \mathbf{v} und \mathbf{w} sind also Linearkombinationen von \mathbf{v}' und \mathbf{w}' . Und umgekehrt. Mit der Notation

Definition 1.15. Sind \mathbf{v} und \mathbf{w} zwei linear unabhängige Vektoren, so heißt

$$\langle \mathbf{v}, \mathbf{w} \rangle := \{ \lambda \mathbf{v} + \mu \mathbf{w} : \lambda, \mu \in \mathbb{R} \}$$

Erzeugnis von \mathbf{v} und \mathbf{w} .

ergibt sich

Satz 1.16. Die zwei Ebenen \mathcal{E} und \mathcal{E}' sind genau dann parallel, wenn die Erzeugnisse der Richtungs-paare gleich sind, d.h., wenn gilt

$$\langle \mathbf{v}, \mathbf{w} \rangle = \langle \mathbf{v}', \mathbf{w}' \rangle .$$

Bemerkung: Für linear unabhängige Vektoren \mathbf{v} und \mathbf{w} ist das Erzeugnis $V := \langle \mathbf{v}, \mathbf{w} \rangle$ selbst wieder eine Ebene — sie geht durch den Ursprung. V erfüllt alle Bedingungen an einen Vektorraum (u.a.: die Summe von zwei Vektoren aus V und das Vielfache eines Vektors aus V liegt wieder in V) — man erhält einen *Untervektorraum* von \mathbb{R}^n der Dimension 2 mit Basis \mathbf{v} und \mathbf{w} .

1.8.2 Komplanarität

Vier Punkte P, Q, R und S heißen **komplanar**, wenn sie in einer Ebene liegen.

Wir betrachten die beiden Vektoren $\mathbf{v} := \vec{PQ}$, $\mathbf{w} := \vec{PR}$. Wenn \mathbf{v} und \mathbf{w} linear abhängig sind, so sind die drei Punkte P, Q und R kollinear (Satz 1.14) und die vier Punkte P, Q, R und S liegen auf jeden Fall in einer Ebene (eventuell sogar auf einer Geraden!). Seien also \mathbf{v} und \mathbf{w} linear unabhängig. Dann kommt als gemeinsame Ebene aller vier Punkte nur die P, Q und R enthaltende Ebene

$$\mathcal{E} := \{ P + \lambda \mathbf{v} + \mu \mathbf{w} : \lambda, \mu \in \mathbb{R} \}$$

in Frage²⁹. \mathcal{E} ist die durch die drei nicht-kollinearen Punkte P, Q und R eindeutig bestimmte Ebene, die diese drei Punkte enthält. Nun betrachte man als weiteren Vektor $\mathbf{u} := \vec{PS}$. Wenn $S \in \mathcal{E}$, muss \mathbf{u} eine Linearkombination von \mathbf{v} und \mathbf{w} sein, also im Erzeugnis $\langle \mathbf{v}, \mathbf{w} \rangle$ liegen. Die drei Vektoren \mathbf{v}, \mathbf{w} und \mathbf{u} bilden dann ein linear abhängiges System. Wir erhalten

Satz 1.17. Die vier Punkte P, Q, R und S sind komplanar genau dann, wenn die drei Vektoren \vec{PQ}, \vec{PR} und \vec{PS} ein linear abhängiges System bilden.

²⁹Dass $Q \in \mathcal{E}$ liegt an $Q = P + 1 \cdot \mathbf{v} + 0 \cdot \mathbf{w}$. Entsprechend erhält man R durch $\lambda = 0$ und $\mu = 1$.

Im \mathbb{R}^3 ist also Nicht-Komplanarität durch die lineare Unabhängigkeit eines Systems aus drei Vektoren garantiert, die wir wie in (1.1) zu einer 3×3 -Matrix zusammenstellen können. Später werden wir sehen, dass es wieder eine Determinantenbedingung gibt, die die lineare Unabhängigkeit gewährleistet.

In der Formulierung von Satz 1.17 wird der Punkt P „bevorzugt“. Aber dies ist nur eine scheinbare Bevorzugung. So sind die beiden Aussagen „Die drei Vektoren \vec{PQ} , \vec{PR} und \vec{PS} bilden ein linear abhängiges System“ und „Die drei Vektoren \vec{QP} , \vec{QR} und \vec{QS} bilden ein linear abhängiges System“ gleichwertig.

Beispiel: Gegeben seien die vier Punkte $P = (0, 0, 0)$, $Q = (0, 2, 1)$, $R = (1, 0, 0)$, $S = (3, 2, 1)$. Sind diese komplanar? Berechne die drei Richtungsvektoren $\mathbf{v}_1 = \vec{PQ} = (0, 2, 1)$, $\mathbf{v}_2 = \vec{PR} = (1, 0, 0)$ und $\mathbf{v}_3 = \vec{PS} = (3, 2, 1)$. Zur Untersuchung der linearen Unabhängigkeit des aus den drei Vektoren $\mathbf{v}_j, j = 1, 2, 3$ bestehenden Systems gehen wir von einer Linearkombination

$$\mathbf{0} = c_1 \mathbf{v}_1 + c_2 \mathbf{v}_2 + c_3 \mathbf{v}_3,$$

aus. Komponentenweise aufgeschrieben gilt also

$$0 = c_1 \cdot 0 + c_2 \cdot 1 + c_3 \cdot 3,$$

$$0 = c_1 \cdot 2 + c_2 \cdot 0 + c_3 \cdot 2,$$

$$0 = c_1 \cdot 1 + c_2 \cdot 0 + c_3 \cdot 1.$$

Dies ist ein lineares Gleichungssystem für c_1, c_2, c_3 , das sich zu $0 = c_2 + 3c_3$, $0 = 2c_1 + 2c_3$, $0 = c_1 + c_3$ vereinfacht. Dies besitzt eine „nichttriviale“ Lösung $c_1 = 1$, $c_3 = -1$, $c_2 = 3$. Die 4 Punkte sind komplanar. Dies hätte man schneller sehen können, wenn man $\mathbf{v}_3 = 3\mathbf{v}_2 + \mathbf{v}_1$ bemerkt hätte.

1.8.3 Schnitt von Ebene und Gerade

Sei $\mathcal{E} = \{P + \lambda \mathbf{v} + \mu \mathbf{w} : \lambda, \mu \in \mathbb{R}\}$ eine Ebene und $\mathcal{G} = \{P' + \lambda' \mathbf{v}' : \lambda' \in \mathbb{R}\}$ eine Gerade im \mathbb{R}^3 . Beide können als parallel zu einander angesehen werden, wenn $\mathbf{v}' \in \langle \mathbf{v}, \mathbf{w} \rangle$. Denn dann gibt es eine zu \mathcal{G} parallele Gerade in \mathcal{E} .

Wenn also \mathcal{E} und \mathcal{G} zueinander parallel sind, ist es denkbar (wenn auch „unwahrscheinlich³⁰“), dass $\mathcal{G} \subset \mathcal{E}$, d.h. dass die Gerade \mathcal{G} in der Ebene \mathcal{E} verläuft. Im letzteren Fall ist die Schnittmenge $\mathcal{E} \cap \mathcal{G} = \mathcal{G}$. Ansonsten gibt es keinen Schnittpunkt.

Wir nehmen daher an, dass keine Parallelität von \mathcal{G} und \mathcal{E} vorliegt. Mathematisch bedeutet dies:

Das System, bestehend aus Richtungspaar \mathbf{v}, \mathbf{w} der Ebene und Richtungsvektor \mathbf{v}' der Gerade, seien linear unabhängig, d.h. sie bilden eine Basis des \mathbb{R}^3 . Wir erwarten auf Grund unserer geometrischen Intuition, dass es dann genau einen Schnittpunkt $S \in \mathcal{E} \cap \mathcal{G}$ gibt. Das wollen wir

³⁰bei „zufälliger“ Auswahl von \mathcal{E} und \mathcal{G}

jetzt auch zeigen und gleichzeitig einen Algorithmus andeuten, wie man diesen Schnittpunkt berechnet. Wichtig ist für mich aber nur, dass Sie verstehen, dass die Suche nach S durch ein *lineares Gleichungssystem* mit drei Unbekannten und drei Gleichungen gegeben ist, welches genau eine Lösung hat.

Der Schnittpunkt S muss sich einerseits mittels λ und μ als Punkt der Ebene \mathcal{E} durch $S = P + \lambda\mathbf{v} + \mu\mathbf{w}$ und andererseits mittels λ' als Punkt von \mathcal{G} durch $S = P' + \lambda'\mathbf{v}'$ darstellen lassen. Gesucht sind also die drei reellen Zahlen λ, μ und λ' mit

$$P + \lambda\mathbf{v} + \mu\mathbf{w} = P' + \lambda'\mathbf{v}'. \quad (1.11)$$

Da es jeweils drei Komponenten der Punkte und Vektoren gibt, handelt es sich um drei lineare Gleichungen. Nun sieht man auch rasch, warum die obige Basis-Annahme an das System $\mathbf{v}, \mathbf{w}, \mathbf{v}'$ zu einer *eindeutigen Lösung* führt: Umformung von (1.11) ergibt

$$\lambda\mathbf{v} + \mu\mathbf{w} - \lambda'\mathbf{v}' = P' - P.$$

Die rechte Seite $\mathbf{a} := P' - P$ ist ein Vektor aus \mathbb{R}^3 und besitzt wegen der Basis-Eigenschaft eine eindeutige Linearkombination durch die drei Vektoren \mathbf{v}, \mathbf{w} und \mathbf{v}' . Die Linearkoeffizienten liefern direkt die drei gesuchten Parameter (nur bei dem dritten muss man das Vorzeichen umkehren). Der gesuchte Schnittpunkt ist dann $S = P' + \lambda'\mathbf{v}'$.

1.8.4 Schnittgerade zweier Ebenen

Wir erwarten, dass sich in \mathbb{R}^3 zwei nicht-parallele Ebenen in einer gemeinsamen Geraden schneiden. Wie findet man diese Schnittgerade?

Die Idee zur Lösung ist einfach: Man nehme zwei parallele Geraden der ersten Ebene und schneide sie mit der zweiten Ebene, wie im vorherigen Kap. 1.8.3 vorgemacht. Sodann bestimme man die Verbindungsgerade der beiden Schnittpunkte. Das ist kein Problem: Sind S_1 und S_2 die Schnittpunkte, so ist $\vec{S_1S_2}$ Richtungsvektor und $P := S_1$ ein Punkt auf der Geraden.

Bleibt die Frage, wie man in

$$\mathcal{E} = \{P + \lambda\mathbf{v} + \mu\mathbf{w} : \lambda, \mu \in \mathbb{R}\}$$

zwei parallele Geraden bestimmt, die sich mit der Ebene

$$\mathcal{E}' = \{P' + \lambda\mathbf{v}' + \mu\mathbf{w}' : \lambda, \mu \in \mathbb{R}\}$$

in jeweils genau einem Punkt schneiden: Wir haben angenommen, dass die beiden Ebenen nicht parallel sind, d.h. dass \mathbf{v} oder \mathbf{w} nicht im Erzeugnis $\langle \mathbf{v}', \mathbf{w}' \rangle$ liegt. Dann ist das aus $\mathbf{v}, \mathbf{v}', \mathbf{w}'$ oder³¹ das aus $\mathbf{w}, \mathbf{v}', \mathbf{w}'$ bestehende System linear unabhängig, liefert also eine Basis des \mathbb{R}^3 .

³¹kein ausschließendes oder!

Wir nehmen ersteres an. Dann nehmen wir zwei (parallele) Geraden in \mathcal{E} mit Richtungsvektor \mathbf{v} , z.B.

$$\mathcal{G}_1 = \{P + \lambda \mathbf{v} : \lambda \in \mathbb{R}\}$$

und

$$\mathcal{G}_2 = \{P + \mathbf{w} + \lambda \mathbf{v} : \lambda \in \mathbb{R}\}.$$

(Setzt man im ersten Fall $\mu = 0$ und im zweiten Fall $\mu = 1$ für einen allgemeinen Punkt $P + \lambda \mathbf{v} + \mu \mathbf{w}$ der Ebene \mathcal{E} , so sieht man, dass die beiden Geraden \mathcal{G}_1 und \mathcal{G}_2 in \mathcal{E} liegen). So, und jetzt wende man das Vorgehen aus dem letzten Abschnitt 1.8.3 zweimal an!

1.8.5 Koordinatenform von Ebenen

Wie für Geraden in der Ebene (Kap. 1.7.1) gibt es auch für Ebenen im Raum eine alternative Beschreibung durch *Gleichungen* an Stelle der Parameterform. Die Gleichung ist eine lineare Gleichung in den drei Komponenten, nämlich

$$ax_1 + bx_2 + cx_3 = d,$$

genauer: Die Punktmenge

$$\mathcal{E} := \{(x_1, x_2, x_3) \in \mathbb{R}^3 : ax_1 + bx_2 + cx_3 = d\}$$

zu gegebenen Koeffizienten a, b, c, d ist in unserem Blick. Wie bei den Geraden sollten wir annehmen, dass nicht alle drei Koeffizienten a, b und c gleichzeitig Null sind. Nehmen wir einmal an, dass $c \neq 0$. Dann können wir nach der „Höhe“ x_3 auflösen und erhalten

$$\mathcal{E} = \{(x_1, x_2, x_3) \in \mathbb{R}^3 : x_3 = a'x_1 + b'x_2 + d'\}$$

mit $a' = -\frac{a}{c}$, $b' = -\frac{b}{c}$, $d' = \frac{d}{c}$. \mathcal{E} ist also der Graph einer Funktion f in zwei Variablen, nämlich $f : (x_1, x_2) \mapsto a'x_1 + b'x_2 + d'$. Irgendwie spürt man, dass \mathcal{E} wirklich eine Ebene ist. Aber können wir sie in eine Parameterform überführen? Das ist einfacher, als es aussieht. Man setze einfach $\lambda = x_1$ und $\mu = x_2$ als Parameter und erhält die Parameterdarstellung

$$x_1 = \lambda, \quad x_2 = \mu, \quad x_3 = a'\lambda + b'\mu + d'$$

für einen Punkt von \mathcal{E} . Damit ist wirklich

$$\mathcal{E} = \{P + \lambda \mathbf{v} + \mu \mathbf{w} : \lambda, \mu \in \mathbb{R}\}$$

mit $P = (0, 0, d')$, $\mathbf{v} = (1, 0, a')$, $\mathbf{w} = (0, 1, b')$.

Wenn wir das Skalarprodukt kennen, wissen wir noch mehr über die geometrische Bedeutung des Vektors (a, b, c) für eine Ebene in Koordinatenform

$$\mathcal{E} := \{(x_1, x_2, x_3) \in \mathbb{R}^3 : ax_1 + bx_2 + cx_3 = d\}.$$

Es wird sich herausstellen, dass dieser Vektor senkrecht auf \mathcal{E} steht.

Beispiel: Gegeben seien die drei Punkte $P = (0, -2, 1)$, $Q = (1, 0, 2)$, $R = (0, 0, 1)$. Die zugehörige Ebene wird durch

$$x_1 = \lambda, \quad x_2 = -2 + 2\lambda + 2\mu, \quad x_3 = 1 + \lambda$$

parametrisiert, siehe (1.10). Nun eliminieren wir $\lambda = x_1$ (und $\mu = (x_2 + 2 - 2\lambda)/2 = (x_2 + 2 - 2x_1)/2$, was aber nicht nötig tut) und setzen dies in die letzte Gleichung ein: Wir erhalten die Ebenengleichung

$$x_3 = 1 + x_1$$

oder $x_1 - x_3 = -1$ bzw.

$$\mathcal{E} := \{(x_1, x_2, x_3) \in \mathbb{R}^3 : x_1 - x_3 = -1\}.$$

Kapitel 2

Skalarprodukt, Winkel, Orthogonalität

2.1 Einführung

Im Anschauungsraum \mathbb{R}^2 und \mathbb{R}^3 kann man offensichtlich von *Winkeln* zwischen zwei Vektoren sprechen. Hier stehen zwei verschiedene Möglichkeiten (γ_1 oder γ_2 in Abb. 2.1) zur Auswahl, wobei sich die beiden Winkel zu 360 Grad (dies entspricht 2π im Bogenmaß) aufaddieren. Man hat sich dafür entschieden, den kleineren der beiden Winkel zu nehmen. Im Folgenden sind alle Winkel zwischen 0 Grad und 180 Grad, bzw. im Bogenmaß (für das wir uns entscheiden) zwischen 0 und π .

Wir haben bisher von dem Winkelbegriff nur im Zusammenhang mit den „kartesischen“ Koordinatensystemen mit seinen senkrecht aufeinander stehenden Koordinatenachsen Gebrauch gemacht — wenn z.B. die auf den Satz von Pythagoras beruhende Formel

$$|\mathbf{x}| = \sqrt{\sum_{j=1}^n x_j^2}$$

für die Länge eines Vektors mit unserer Vorstellung übereinstimmen soll.

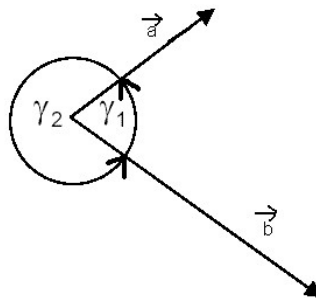


Abbildung 2.1: Winkel zwischen Vektoren

Dass wir das Wort *Skalarprodukt* in der Benennung dieses Kapitels noch *vor* dem Wort *Winkel* platziert haben, liegt daran, dass praktisch nie mit dem Winkel zwischen den Vektoren \mathbf{v} und \mathbf{w} , genannt, $\angle(\mathbf{v}, \mathbf{w})$, gerechnet wird. Stattdessen taucht der Term $\mathbf{v} \cdot \mathbf{w}$ auf — dies ist eine *reelle Zahl*, eben das Skalarprodukt von \mathbf{v} und \mathbf{w} . Hier ist „ \cdot “ *keine* binäre Verknüpfung: Das „Produkt“ der beiden Vektoren ist kein Vektor, sondern ein „Skalar“ (eine Zahl) — deshalb eben *Skalarprodukt*¹. Es gelten jedoch ganz ähnliche Rechenregeln wie beim „normalen“ Produkt. Nicht verwechselt werden sollte das Skalarprodukt mit dem Produkt zwischen einem Skalar λ und einem Vektor \mathbf{x} !

Besonders wichtig ist es, dass $\mathbf{v} \cdot \mathbf{w} = 0$ genau dann gilt, wenn die Vektoren \mathbf{v} und \mathbf{w} *senkrecht* (*orthogonal*) zueinander sind, wenn also $\angle(\mathbf{v}, \mathbf{w}) = \frac{\pi}{2}$.

Als schöne Anwendungen dieses Konzepts lassen sich die klassischen Sätze von Pythagoras und von Thales leicht „beweisen“, wobei man aber beachten muss, dass der Satz von Pythagoras schon benutzt wird, um die Definition des Winkels über das Skalarprodukt mit unserer Anschauung übereinstimmen zu lassen. Auch die *Hesseschen Normalformen* von Geraden und Ebenen gründen sich auf Skalarprodukte. Die *Koordinatenformen* von Geraden und Ebenen in Kap. 1.7.1 und 1.8.5 erscheinen nun in einem anderen Licht.

Erst unter Einbeziehung des Skalarprodukts wird der \mathbb{R}^n zu einem *euklidischen Vektorraum*. Die Bedeutung des Skalarprodukts (und allgemein von euklidischen Vektorräumen) geht aber noch viel weiter und kann im Rahmen dieser Einführung nicht vertieft werden. Ich erwähne nur als ein Beispiel die *Fourier-Analyse* mit trigonometrischen Funktionen.

2.2 Skalarprodukt von zwei Vektoren und Winkel zwischen Vektoren

Definition 2.1. Gegeben seien zwei Vektoren $\mathbf{a} = (a_1, \dots, a_n)$, $\mathbf{b} = (b_1, \dots, b_n) \in \mathbb{R}^n$. Das **Skalarprodukt** von \mathbf{a} und \mathbf{b} ist durch

$$\mathbf{a} \cdot \mathbf{b} := a_1 b_1 + a_2 b_2 + \dots + a_n b_n,$$

kurz durch

$$\mathbf{a} \cdot \mathbf{b} := \sum_{j=1}^n a_j b_j$$

definiert.

Bemerkung: Dass wir das für „normale“ Multiplikation reeller Zahlen vorgesehene Malzeichen „ \cdot “ auch hier einsetzen, kann zu Missverständnissen führen. Bedenken Sie aber, dass für $n = 1$ die „Vektoren“ Skalare sind und dass dann das Skalarprodukt zu einem normalen Produkt wird.

¹Das ist beim *Vektorprodukt* in Kap. 2.6 anders.

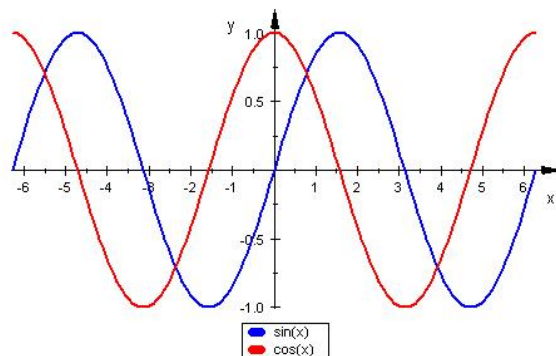


Abbildung 2.2: Graphen der Sinus- und Kosinus-Funktion

Auf jeden Fall muss man sich immer klar sein, was man mit „ \cdot “ gerade meint. Ist z.B. $\lambda \in \mathbb{R}$, so involviert die Rechenregel $(\lambda \cdot \mathbf{v}) \cdot \mathbf{w} = \lambda \cdot (\mathbf{v} \cdot \mathbf{w})$ drei verschiedene Produkte (Skalar mal Skalar, Skalar mal Vektor und Vektor mal Vektor).

Die Bedeutung des Skalarprodukts liegt darin, dass man mit seiner Hilfe den *Winkel* zwischen zwei Vektoren berechnen kann. Dazu muss man die *Kosinus*-Funktion kennen und wissen, dass zu gegebenem $y \in \mathbb{R}$ mit $-1 \leq y \leq 1$ der Winkel ϕ durch die Gleichung $\cos \phi = y$ eindeutig festgelegt ist, wenn man sich (im Bogenmaß!) auf $0 \leq \phi \leq \pi$ festlegt. Dass dies so ist, liegt daran, dass die Kosinus-Funktion² eine Bijektion zwischen $[0, \pi]$ und $[-1, 1]$ ist, s. Abb. 2.2.

In der Ebene meinen wir zu wissen, was wir unter dem Winkel zwischen zwei Vektoren verstehen. Abb. 2.1 zeigt zwei Vektoren $\vec{\mathbf{a}}$ und $\vec{\mathbf{b}}$ und die zwei von ihnen eingeschlossenen Winkel γ_1 und γ_2 . Der Winkel, den wir meinen, ist stets der kleinere der beiden dort angezeigten Winkel, also der Winkel γ_1 , der im Bogenmaß stets zwischen 0 und π (und im Gradmaß zwischen 0 und 180 Grad) liegt. Im Raum haben wir dieselbe Situation, wir müssen die Abbildung 2.1 nur so interpretieren, dass die Zeichenebene die Ebene ist, in der die beiden Vektoren liegen.

Wie hängen nun aber Skalarprodukt und Winkel zusammen? Zunächst einmal stellen wir fest, dass der Winkel zwischen zwei Vektoren nur definiert ist, wenn beide Vektoren eine positive Länge haben, also wenn keiner der beiden Vektoren der Nullvektor ist. Dann darf der Winkel auch nicht von der Länge der Vektoren abhängen. Daher nehmen wir zunächst an, dass die Vektoren normiert sind, d.h. dass $|\mathbf{a}| = 1$ und $|\mathbf{b}| = 1$ gelten. Dann definieren wir

Definition 2.2. Seien \mathbf{a} und \mathbf{b} zwei Einheitsvektoren des \mathbb{R}^n , d.h. zwei Vektoren der Länge Eins. Dann ist der Winkel³ $\angle(\mathbf{a}, \mathbf{b})$ zwischen \mathbf{a} und \mathbf{b} durch

$$\cos \angle(\mathbf{a}, \mathbf{b}) := \mathbf{a} \cdot \mathbf{b}$$

definiert.

²Wenn Ihre Kenntnisse über trigonometrische Funktionen aufgefrischt werden sollen, schauen Sie unter [Winkelfunktionen](#) (Mathe Online Wien) nach.

³Eigentlich nur der Kosinus des Winkels

Diese Definition gibt nur dann einen Sinn, wenn $|\mathbf{a} \cdot \mathbf{b}| \leq 1$ gilt, da ja für alle Winkel ϕ gilt: $|\cos \phi| \leq 1$. Dass dem so ist, liegt an

Satz 2.3 (Cauchy-Schwarzsche Ungleichung). *Für alle Vektoren $\mathbf{a}, \mathbf{b} \in \mathbb{R}^n$ gilt*

$$|\mathbf{a} \cdot \mathbf{b}| \leq |\mathbf{a}| \cdot |\mathbf{b}|.$$

Denn wenn wir in dieser Ungleichung $|\mathbf{a}| = 1$ und $|\mathbf{b}| = 1$ einsetzen, erhalten wir die gewünschte Ungleichung $|\mathbf{a} \cdot \mathbf{b}| \leq 1$. Wir stellen den Beweis zu diesem Satz zurück und begnügen uns damit zu erklären, warum die Def. 2.2 mit der Abb. 2.1 verträglich ist. Dazu denken wir uns ein kartesisches Koordinatensystem in die Abb. 2.1 gelegt mit \mathbf{b} in Richtung der x -Achse. Wenn $|\mathbf{b}| = 1$, ist $\mathbf{b} = (1, 0)$. Es ist dann $\mathbf{a} \cdot \mathbf{b} = a_1$ und nach Definition des Kosinus⁴ gilt $\cos(\angle(\mathbf{a}, \mathbf{b})) = \frac{a_1}{|\mathbf{a}|}$. Wenn nun auch $|\mathbf{a}| = 1$, haben wir Def. 2.2 „bestätigt“.

Man könnte im \mathbb{R}^2 und im \mathbb{R}^3 auch davon ausgehen, dass der Begriff *Winkel* $\angle(\mathbf{a}, \mathbf{b})$ ein wohl definierter Begriff ist und den Sachverhalt

$$\cos \angle(\mathbf{a}, \mathbf{b}) = \mathbf{a} \cdot \mathbf{b}$$

für $|\mathbf{a}| = 1$ und $|\mathbf{b}| = 1$ *beweisen*. Das haben wir eben fast auch getan, wenn man noch rechtfertigt, dass man das kartesische Koordinatensystem einfach so wählen kann, wie hier geschehen. Dazu müssten wir zeigen, dass $\mathbf{a} \cdot \mathbf{b}$ in der hier definierten *koordinatenabhängigen* Form

$$\mathbf{a} \cdot \mathbf{b} := a_1 b_1 + a_2 b_2 + \cdots + a_n b_n$$

sich nicht ändert, wenn wir ein anderes (kartesisches!) Koordinatensystem wählen. Das unterlassen wir aber hier, werden das aber in Kap. 2.3.3 nachholen.

Man kann aber auch den Kosinussatz benutzen, um diesen Sachverhalt zu *beweisen*. Dazu betrachten wir ein Dreieck mit den Seitenlängen $|\mathbf{a}|$ und $|\mathbf{b}|$ und dem Winkel $\gamma := \angle(\mathbf{a}, \mathbf{b})$. Der Kosinussatz besagt

$$|\mathbf{b} - \mathbf{a}|^2 = |\mathbf{a}|^2 + |\mathbf{b}|^2 - 2|\mathbf{a}||\mathbf{b}|\cos(\gamma).$$

Es ist $|\mathbf{b} - \mathbf{a}|^2 = (\mathbf{b} - \mathbf{a}) \cdot (\mathbf{b} - \mathbf{a}) = |\mathbf{a}|^2 + |\mathbf{b}|^2 - 2\mathbf{a} \cdot \mathbf{b}$, so dass folgt

$$\mathbf{a} \cdot \mathbf{b} = |\mathbf{a}||\mathbf{b}|\cos(\gamma).$$

Jetzt kann die Def. 2.2 verallgemeinert werden:

Definition 2.4. *Seien $\mathbf{a} \neq \mathbf{o}$ und $\mathbf{b} \neq \mathbf{o}$. Dann setzt man*

$$\cos(\angle(\mathbf{a}, \mathbf{b})) := \frac{\mathbf{a} \cdot \mathbf{b}}{|\mathbf{a}||\mathbf{b}|} \tag{2.1}$$

⁴Ankathete durch Hypotenuse

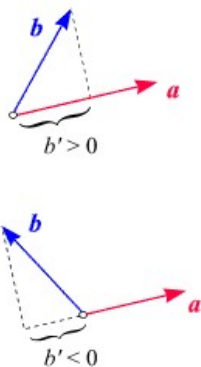


Abbildung 2.3: Winkel zwischen Vektoren

Man ersetze nur in Def. 2.2 \mathbf{a} durch seine „Normierung“ $\frac{\mathbf{a}}{|\mathbf{a}|}$ und \mathbf{b} durch $\frac{\mathbf{b}}{|\mathbf{b}|}$.

Es hat sich gezeigt, dass es gar nicht so günstig ist, mit dem Winkel zwischen zwei Vektoren zu arbeiten. Einen vollständigen Ersatz liefert das Skalarprodukt $\mathbf{a} \cdot \mathbf{b}$! Schauen wir uns Abb. 2.3 an, so sehen wir, dass die Größe b' entscheidend ist! Diese ist aber gerade durch

$$|\mathbf{a}| \cdot b' = \mathbf{a} \cdot \mathbf{b} \quad (2.2)$$

definiert, was man sofort einsieht, wenn man das Koordinatensystem so legt, dass $\mathbf{a} = (a_1, 0)$. Während $\angle(\mathbf{v}, \mathbf{w})$ keinen Sinn macht, wenn $\mathbf{v} = \mathbf{o}$ oder $\mathbf{w} = \mathbf{o}$, ist dies beim Skalarprodukt anders. Hier gilt für alle Vektoren \mathbf{v} und \mathbf{w}

$$\mathbf{o} \cdot \mathbf{w} = 0 = \mathbf{v} \cdot \mathbf{o}.$$

2.2.1 Rechenregeln für das Skalarprodukt

In einigen Lehrbüchern werden Skalarprodukte in Vektorräumen definiert, ohne sich wie hier in Def. 2.1 auf den \mathbb{R}^n zu beziehen⁵. Stattdessen werden Eigenschaften angegeben, die in unserem Fall

$$\mathbf{a} \cdot \mathbf{b} := \sum_{j=1}^n a_j b_j$$

offensichtlich gültige „Rechenregeln“ sind:

Definitheit: Für alle Vektoren \mathbf{v} gilt

$$\mathbf{v} \cdot \mathbf{v} \geq 0.$$

Es gilt $\mathbf{v} \cdot \mathbf{v} = 0$ genau dann, wenn $\mathbf{v} = \mathbf{o}$.

⁵Hier wurde das Skalarprodukt unter Zugrundelegung eines kartesischen Koordinatensystems definiert, das es erlaubt, von Komponenten der Vektoren \mathbf{a} und \mathbf{b} zu sprechen. Es geht auch anders — etwa so, wie wir es für das Vektorprodukt in Kap. 2.6 tun werden.

Homogenität: Für alle Skalare $\lambda \in \mathbb{R}$ und alle Vektoren \mathbf{u}, \mathbf{v} gilt

$$(\lambda \cdot \mathbf{u}) \cdot \mathbf{v} = \lambda \cdot (\mathbf{u} \cdot \mathbf{v}).$$

Distributivität: Für alle Vektoren $\mathbf{u}, \mathbf{v}, \mathbf{w}$ gilt

$$(\mathbf{u} + \mathbf{v}) \cdot \mathbf{w} = \mathbf{u} \cdot \mathbf{w} + \mathbf{v} \cdot \mathbf{w}.$$

Symmetrie, Kommutativität: Für alle Vektoren \mathbf{u}, \mathbf{v} gilt

$$\mathbf{u} \cdot \mathbf{v} = \mathbf{v} \cdot \mathbf{u}.$$

Die Rechenregel

$$(\mathbf{u} + \mathbf{v}) \cdot \mathbf{w} = \mathbf{u} \cdot \mathbf{w} + \mathbf{v} \cdot \mathbf{w}$$

ähnelte auf dem ersten Blick einem Distributivgesetz, die Regel

$$(\lambda \cdot \mathbf{u}) \cdot \mathbf{v} = \lambda \cdot (\mathbf{u} \cdot \mathbf{v})$$

einem Assoziativgesetz und die Symmetrie einem Kommutativgesetz. Klar, dass auch noch die Regeln

$$\mathbf{u} \cdot (\lambda \cdot \mathbf{v}) = \lambda \cdot (\mathbf{u} \cdot \mathbf{v})$$

und

$$\mathbf{w} \cdot (\mathbf{u} + \mathbf{v}) = \mathbf{w} \cdot \mathbf{u} + \mathbf{w} \cdot \mathbf{v}$$

gelten.

Insbesondere gilt $(-\mathbf{v}) \cdot \mathbf{w} = -\mathbf{v} \cdot \mathbf{w}$, woraus

$$\sphericalangle(-\mathbf{v}, \mathbf{w}) = \pi - \sphericalangle(\mathbf{v}, \mathbf{w})$$

folgt, s. Übungen und Abb. 2.4.

2.2.2 Länge eines Vektors und Skalarprodukt

Es wurde schon darauf hingewiesen, dass die Länge eines Vektors (s. Kap. 1.6),

$$|\mathbf{x}| = \sqrt{x_1^2 + \cdots + x_n^2}$$

entscheidend mit (rechten) Winkeln zu tun hat. Natürlich auch mit dem Skalarprodukt, weil

$$|\mathbf{x}| = \sqrt{\mathbf{x} \cdot \mathbf{x}}$$

bzw.

$$|\mathbf{x}|^2 = \mathbf{x} \cdot \mathbf{x}.$$

Insofern können wir die Längen $|\mathbf{a}|$ und $|\mathbf{b}|$ vollständig aus der Winkel-Formel (2.1) verbannen und erhalten

$$\cos(\sphericalangle(\mathbf{a}, \mathbf{b})) = \frac{\mathbf{a} \cdot \mathbf{b}}{\sqrt{(\mathbf{a} \cdot \mathbf{a})(\mathbf{b} \cdot \mathbf{b})}}.$$

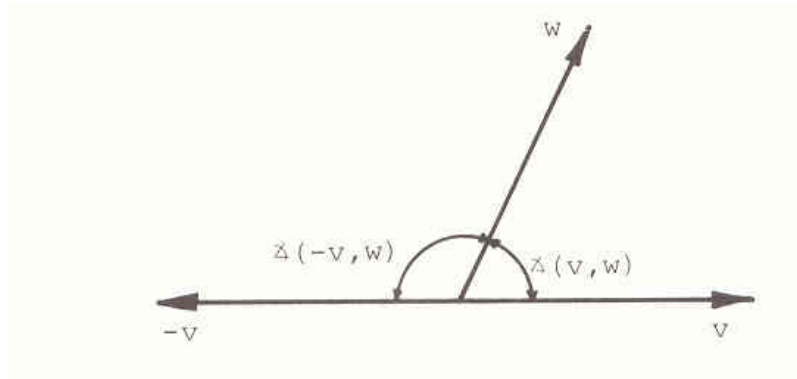


Abbildung 2.4: $\angle(-\mathbf{v}, \mathbf{w}) = \pi - \angle(\mathbf{v}, \mathbf{w})$

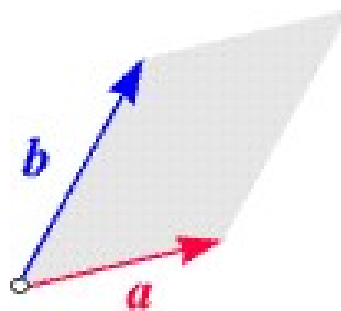


Abbildung 2.5: Fläche eines Parallelogramms

2.2.3 Formel für die Fläche eines Parallelogramms

Ein Parallelogramm ist durch seine Kantenvektoren \mathbf{a} und \mathbf{b} festgelegt, s. Abb. 2.5.

Man kann dessen Fläche A durch die Formel „Länge der Grundlinie \times Höhe“, also durch $A = h \cdot |\mathbf{a}|$ berechnen ⁶. Nun gilt $h^2 + b'^2 = |\mathbf{b}|^2$, s. Abb. 2.6.

b' kann man mittels des Skalarprodukts berechnen, siehe (2.2): Danach gilt $b' \cdot |\mathbf{a}| = \mathbf{a} \cdot \mathbf{b}$, wobei

⁶Diese in der Mittelstufe behandelte Formel kann man ganz einfach aus der für eine Rechtecksfläche herleiten. Man verlege nur das linke Dreieck in Abb. 2.5 ans rechte Ende.

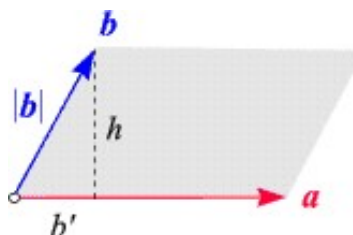


Abbildung 2.6:

$b' > 0$ (bzw. < 0) für einen spitzen (bzw. stumpfen) Winkel $\angle(\mathbf{a}, \mathbf{b})$ gilt. Damit berechnet sich $A^2 = (|\mathbf{b}|^2 - b'^2)|\mathbf{a}|^2 = |\mathbf{a}|^2|\mathbf{b}|^2 - (\mathbf{a} \cdot \mathbf{b})^2$, und wir haben den

Satz 2.5. Für die Fläche A eines von den Kantenvektoren \mathbf{a} und \mathbf{b} gebildeten Parallelogramms gilt die Formel

$$A^2 = |\mathbf{a}|^2|\mathbf{b}|^2 - (\mathbf{a} \cdot \mathbf{b})^2.$$

Vielleicht erinnern Sie sich an Satz 1.13, der $A = |\Delta(\mathbf{a}, \mathbf{b})|$ aussagte und eine Beziehung zwischen Parallelogrammfläche A und der Determinante $\Delta(\mathbf{a}, \mathbf{b})$ liefert. Dabei ist $\Delta(\mathbf{a}, \mathbf{b}) = a_1b_2 - a_2b_1$, wenn $\mathbf{a} = (a_1, a_2)$ und $\mathbf{b} = (b_1, b_2)$ die beiden Kantenvektoren des Parallelogramms in der Ebene sind. Setzt man letzteres in Satz 2.5 ein, so erhält man die gleiche Formel (s. Übungen).

Jetzt können wir auch wieder den Winkel $\angle(\mathbf{a}, \mathbf{b})$ ins Spiel bringen. Die Fläche A eines Parallelogramms mit Grundkante \mathbf{a} und Höhe h ist ja $A = |\mathbf{a}| \cdot h$. Nun ist $h = |\mathbf{b}| \sin(\angle(\mathbf{a}, \mathbf{b}))$, und es gilt

Satz 2.6. Für die Fläche A eines von den ebenen Kantenvektoren \mathbf{a} und \mathbf{b} gebildeten Parallelogramms gilt die Formel

$$A = |\mathbf{a}||\mathbf{b}| \sin(\angle(\mathbf{a}, \mathbf{b})).$$

Aus Satz 2.6 folgt durch Quadrierung und Anwendung von $\cos^2(\phi) + \sin^2(\phi) = 1$ auch wieder Satz 2.5, wenn man (2.1) benutzt.

Wir fassen die drei Formeln für die Fläche A eines Parallelogramms, ausgedrückt durch dessen Kantenvektoren \mathbf{a} und \mathbf{b} zusammen:

- Satz 1.13

$$A = |\Delta(\mathbf{a}, \mathbf{b})| = |a_1b_2 - a_2b_1|$$

- Satz 2.5

$$A^2 = |\mathbf{a}|^2|\mathbf{b}|^2 - (\mathbf{a} \cdot \mathbf{b})^2$$

- Satz 2.6

$$A = |\mathbf{a}||\mathbf{b}| \sin(\angle(\mathbf{a}, \mathbf{b}))$$

2.2.4 Korrelationen und Skalarprodukt

In der Statistik stellt man sich die Frage, wann man sagen kann, dass zwei Merkmale miteinander korrelieren. Man kennt den Begriff *Korrelationskoeffizient*, der die Korrelation zwischen zwei Merkmalen quantifiziert.

Man stelle sich vor, dass man die beiden Merkmale „Gewicht“ und „Länge“ von Menschen hat. Wenn man diese beiden Merkmale in einer Stichprobe (z.B. den ZuhörerInnen dieser Vorlesung) misst, erhält man bei einer Stichprobe vom Umfang n zwei Vektoren \mathbf{a} und \mathbf{b} aus dem \mathbb{R}^n . Nun bildet man jeweils den Mittelwert und subtrahiert diesen von allen Komponenten. Die Vektoren,

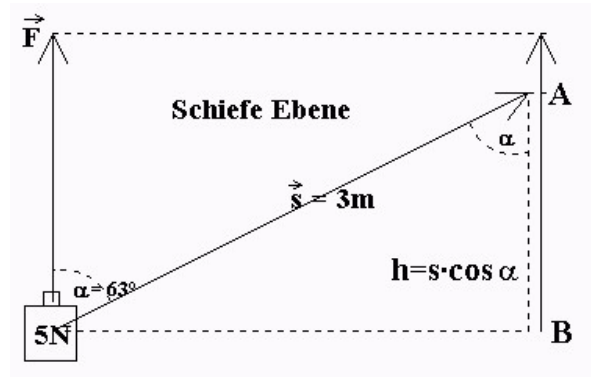


Abbildung 2.7: Schiefe Ebene

die man dann erhält, haben die Zahl 0 als Mittelwert. Dann normiert man sie auf Länge eins, um die Rolle der verwendeten Maßeinheit zu eliminieren. Der gesuchte Korrelationskoeffizient ist dann einfach das Skalarprodukt dieser so normierten Vektoren! Ist dieser Eins, so besteht maximale Korrelation, ist dieser Null, so gibt es keine Korrelation. Ist er gleich -1 , so spricht man von Anti-Korrelation.

Machen Sie einmal einen Versuch: Sie und Ihre FreundIn würfeln $n = 100$ Mal mit einem klassischen Würfel. Dann berechnen Sie den jeweiligen Mittelwert (der in der Nähe von 3,5 liegen sollte), subtrahieren und normieren (Hier brauchen Sie einen Rechner!). Sodann bilden Sie das Skalarprodukt. Wenn das Ergebnis nicht in der Nähe von Null liegt, sind Sie und Ihre FreundIn „würfel-korreliert“!

2.2.5 Ein physikalisches Beispiel

Skalarprodukte von Vektoren spielen auch in der Physik eine Rolle, z.B. ist das Skalarprodukt aus Kraft- und Wegvektor eine Energie (Arbeit). In einem einfachen Fall sieht dies so aus:

Klar dürfte Ihnen sein, dass es anstrengend ist, ein Gewicht hochzuheben. Physikalisch kann man die dazu notwendige *Energie* messen. Einleuchtend ist es, dass diese Energie umso größer ausfällt, je größer das Gewicht ist und je höher Sie es heben sollen. Wenn G das Gewicht ist und h die Höhe, so überrascht es vielleicht nicht, dass hG die aufzubringende Energie ist. Dies unter der Voraussetzung, dass das Gewicht *senkrecht* hochgehoben wird. Was ist nun, wenn man es entlang einer schiefen Ebene wie in Abb. 2.7 transportiert⁷?

Unter Vernachlässigung von Reibung ist auch hier die aufzubringende Energie hG .

Was hat das mit Skalarprodukten zu tun? Nun, das Gewicht (besser: die Masse) wird durch die Gravitation zu einem Vektor, eben \vec{F} , die schiefe Ebene zu einem Vektor \vec{s} . Die aufzubringende Energie ist das Skalarprodukt $\vec{s} \cdot \vec{F}$.

⁷Das Gewicht G heißt hier \vec{F} . Genauer: $G = |\vec{F}|$.

2.3 Orthogonalität

Was bedeutet die Eigenschaft $\mathbf{v} \cdot \mathbf{w} = 0$? Da $\cos(\frac{\pi}{2}) = 0$, ist dies der Fall, wenn die beiden Vektoren einen rechten Winkel einschließen. (Oder natürlich, dass $\mathbf{v} = \mathbf{o}$ oder $\mathbf{w} = \mathbf{o}$). Z.B. gilt $(0, 1) \cdot (-1, 0) = 0$ oder $(1, 1) \cdot (1, -1) = 0$.

Definition 2.7. Zwei Vektoren \mathbf{v} und \mathbf{w} heißen **orthogonal** (in Zeichen $\mathbf{v} \perp \mathbf{w}$; lies: „ \mathbf{v} orthogonal zu \mathbf{w} “), wenn $\mathbf{v} \cdot \mathbf{w} = 0$ gilt.

Aufgabe: Bestimmen Sie einen zu $(3, 1)$ orthogonalen Vektor und zeichnen Sie beide in eine Ebene.

2.3.1 Satz von Pythagoras

Die Katheten eines rechtwinkligen Dreiecks seien durch die Kantenvektoren \mathbf{u} und \mathbf{v} gegeben. Die Hypotenuse ist dann der Vektor $\mathbf{u} + \mathbf{v}$ und der Satz von Pythagoras (570-496 v.Chr) lautet

$$|\mathbf{u}|^2 + |\mathbf{v}|^2 = |\mathbf{u} + \mathbf{v}|^2.$$

Der „Beweis⁸“ ist jetzt einfach, er benutzt natürlich die Orthogonalität $\mathbf{u} \cdot \mathbf{v} = 0$ in dem rechtwinkligen Dreieck mit den Kathetenvektoren \mathbf{u} und \mathbf{v} sowie die vektorielle Eigenschaft, dass $\mathbf{u} + \mathbf{v}$ die Hypotenuse des rechtwinkligen Dreiecks ist:

Es ist $|\mathbf{u} + \mathbf{v}|^2 = (\mathbf{u} + \mathbf{v}) \cdot (\mathbf{u} + \mathbf{v})$. Die Rechenregeln aus Kap. 2.2.1, insbesondere das „Distributivgesetz“ liefert $(\mathbf{u} + \mathbf{v}) \cdot (\mathbf{u} + \mathbf{v}) = \mathbf{u} \cdot \mathbf{u} + \mathbf{u} \cdot \mathbf{v} + \mathbf{v} \cdot \mathbf{u} + \mathbf{v} \cdot \mathbf{v}$. Nun ist $\mathbf{u} \cdot \mathbf{v} = 0 = \mathbf{v} \cdot \mathbf{u}$ sowie $\mathbf{u} \cdot \mathbf{u} = |\mathbf{u}|^2$ und $\mathbf{v} \cdot \mathbf{v} = |\mathbf{v}|^2$. ■

Sie sollten den ursprünglich von Euklid gefundenen Beweis kennen. Einen hervorragenden **Multi-Media-Beweis** (Didaktik der Physik der Uni Würzburg) gibt es im Internet. Das verwendete Java-Applet geht auf J. MOREY zurück.

Es gibt einen einfacheren Beweis, der in Abb. 2.8 wiedergegeben ist, den Sie besser verstehen, wenn Sie die Farben sehen können.

Weiteres zu verschiedenen Beweisen des Satzes von Pythagoras und der mit diesem verwandten Katheten- und Höhensätze finden Sie in **Die Satzgruppe des Pythagoras (Formulierung der Sätze)** (Mathematik-Didaktik der Uni Bayreuth).

In den Übungen werden Sie den Satz von Thales (ca. 650 -560 v.Chr), dass zu jedem Punkt auf einem Halbkreis ein rechter Winkel gemäß Abb. 2.9 konstruiert werden kann⁹, mit Hilfe der Vektorrechnung beweisen dürfen.

⁸Ein Scheinbeweis, da der Satz von Pythagoras schon in der Definition von Länge und Skalarprodukt von Vektoren steckt.

⁹Der elementare „Mittelstufen“-Beweis ist hier auch angedeutet. Man muss nur den Winkelsummensatz geschickt anwenden.

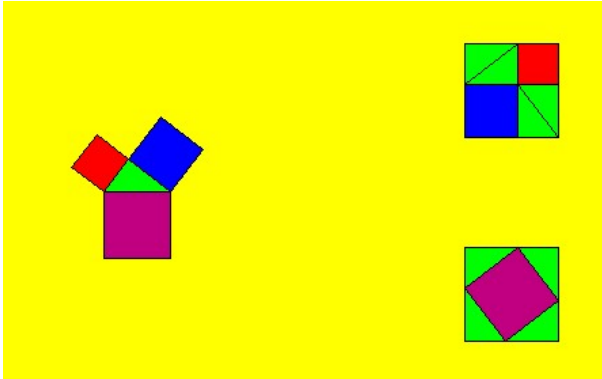


Abbildung 2.8: Beweis des Satzes von Pythagoras

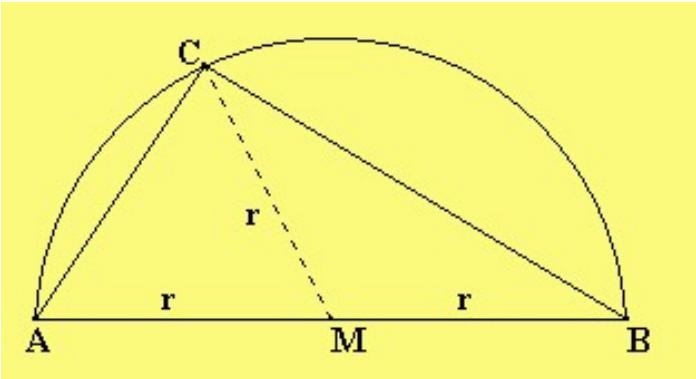


Abbildung 2.9: Zum Satz und Beweis von Thales

2.3.2 Parallelogrammgleichung

Addiert man in einem Parallelogramm alle Quadrate der vier Kantenlängen, so besagt die *Parallelogrammgleichung*, dass diese Summe gleich der Summe der beiden Quadrate der Diagonallängen ist. Für den Spezialfall von Rechtecken ist diese Gleichung mit der Aussage des Satzes von Pythagoras äquivalent. Machen Sie sich dies durch eine Zeichnung klar!

Diesen Sachverhalt können wir ebenfalls einfach mit der Vektorrechnung zeigen: Seien \mathbf{v} und \mathbf{w} die beiden Kantenvektoren des Parallelogramms. Dann sind $\mathbf{u} + \mathbf{v}$ und $\mathbf{u} - \mathbf{v}$ die beiden Diagonalvektoren, und unsere Aussage lautet

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

Wenn man jetzt wie im Beweis des Satzes von Pythagoras $|\mathbf{v} + \mathbf{w}|^2 = (\mathbf{v} + \mathbf{w}) \cdot (\mathbf{v} + \mathbf{w})$ und $|\mathbf{v} - \mathbf{w}|^2 = (\mathbf{v} - \mathbf{w}) \cdot (\mathbf{v} - \mathbf{w})$ ausnutzt und ausmultipliziert sowie die Symmetrie $\mathbf{v} \cdot \mathbf{w} = \mathbf{w} \cdot \mathbf{v}$ einsetzt, erhält man das gewünschte Resultat. ■

2.3.3 Orthogonalbasis

Für die drei Einheitsvektoren $\mathbf{e}_j, j = 1, 2, 3$, des \mathbb{R}^3 gilt, dass je zwei von ihnen orthogonal sind — man sagt, die drei Vektoren sind *paarweise* orthogonal¹⁰. Sie bilden eine *Orthogonalbasis* des \mathbb{R}^3 . Umgekehrt bilden drei von \mathbf{o} verschiedene Vektoren des \mathbb{R}^3 , die paarweise orthogonal sind, stets auch eine Basis des \mathbb{R}^3 . Hierzu ist nach Satz 1.11 nur zu zeigen, dass sie ein linear unabhängiges System bilden.

Satz 2.8. *Seien die drei Vektoren $\mathbf{v}_j, j = 1, 2, 3$, paarweise orthogonal und keiner der Nullvektor. Dann bilden sie ein linear unabhängiges System.*

Der **Beweis** ist nicht schwierig:

Wir untersuchen gemäß Def. 1.7 eine Linearkombination des Nullvektors,

$$\mathbf{o} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + c_3\mathbf{v}_3. \tag{2.3}$$

Multiplizieren wir die linke und rechte Seite dieser Gleichung skalar mit \mathbf{v}_1 , so gilt

$$\mathbf{v}_1 \cdot \mathbf{o} = \mathbf{v}_1 \cdot (c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + c_3\mathbf{v}_3).$$

Nun wenden wir das „Distributivgesetz“ an:

$$\mathbf{v}_1 \cdot \mathbf{o} = \mathbf{v}_1 \cdot (c_1\mathbf{v}_1) + \mathbf{v}_1 \cdot (c_2\mathbf{v}_2) + \mathbf{v}_1 \cdot (c_3\mathbf{v}_3).$$

¹⁰Ist Ihnen wirklich klar, was man unter *n verschiedenen* Zahlen a_1, \dots, a_n versteht? Etwa unter 5 verschiedenen Zahlen? Von 5 *geraden* Zahlen zu sprechen, macht Sinn. Aber von 5 *verschiedenen* Zahlen? Eigentlich ist *verschieden* eine binäre Relation (zwei Dinge sind gleich oder verschieden), und daher ist die Sprechweise „5 *paarweise verschiedene* Zahlen“ korrekt.

Die linke Seite ist Null, mit weiteren Rechenregeln des Skalarprodukts ergibt sich

$$0 = c_1(\mathbf{v}_1 \cdot \mathbf{v}_1) + c_2(\mathbf{v}_1 \cdot \mathbf{v}_2) + c_3(\mathbf{v}_1 \cdot \mathbf{v}_3).$$

Wegen der Orthogonalität gilt $\mathbf{v}_1 \cdot \mathbf{v}_2 = 0$ und $\mathbf{v}_1 \cdot \mathbf{v}_3 = 0$, also

$$0 = c_1(\mathbf{v}_1 \cdot \mathbf{v}_1) = c_1|\mathbf{v}_1|^2,$$

woraus sich wegen $|\mathbf{v}_1| \neq 0$ ergibt, dass der erste Linearkoeffizient $c_1 = 0$. Entsprechend führt die Skalarmultiplikation von (2.3) mit \mathbf{v}_2 und \mathbf{v}_3 zu $c_2 = 0$ und zu $c_3 = 0$. Also gibt es nur die triviale Linearkombination des Nullvektors \mathbf{o} . ■

Definition 2.9. Eine Orthogonalbasis aus $\mathbf{v}_j, j = 1, 2, 3$, des \mathbb{R}^3 heißt eine **Orthonormalbasis**, wenn die \mathbf{v}_j normiert sind, wenn also zusätzlich zur Orthogonalität $|\mathbf{v}_j| = 1, j = 1, 2, 3$, gilt.

Die „Standard“-Orthonormalbasis des \mathbb{R}^3 besteht aus den drei Einheitsvektoren $\mathbf{e}_j, j = 1, 2, 3$. Die Komponenten irgendeines Vektors $\mathbf{x} = (x_1, x_2, x_3)$ ergeben sich aus

$$\mathbf{x} = x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + x_3\mathbf{e}_3.$$

Stellen Sie sich einmal vor, Sie möchten statt des aus \mathbf{e}_j bestehenden Koordinaten-„Dreibeins“ ein anderes kartesisches Koordinatensystem wählen, das durch eine Orthonormalbasis $\mathbf{v}_j, j = 1, 2, 3$ gegeben ist. Die Basis-Eigenschaft liefert eine eindeutige Darstellung eines Vektors \mathbf{x} als Linearkombination $\mathbf{x} = \xi_1\mathbf{v}_1 + \xi_2\mathbf{v}_2 + \xi_3\mathbf{v}_3$ mit Linearkoeffizienten $\xi_j \in \mathbb{R}$. Wie berechnet man die „neuen Komponenten“ ξ_j von \mathbf{x} bzgl. der neuen Basis?

Das geht so: Wir multiplizieren $\mathbf{x} = \xi_1\mathbf{v}_1 + \xi_2\mathbf{v}_2 + \xi_3\mathbf{v}_3$ skalar mit $\mathbf{v}_j, j = 1, 2, 3$ und erhalten wie oben beim Beweis der linearen Unabhängigkeit eines Orthogonalsystems (man nutze $|\mathbf{v}_j| = 1$ aus)

$$\xi_j = \mathbf{x} \cdot \mathbf{v}_j, j = 1, 2, 3.$$

Jetzt können wir auch nachtragen, dass die Berechnungsvorschrift $\mathbf{x} \cdot \mathbf{y} = x_1y_1 + x_2y_2 + x_3y_3$ ganz entsprechend in neuen Koordinaten $\mathbf{x} \cdot \mathbf{y} = \xi_1\eta_1 + \xi_2\eta_2 + \xi_3\eta_3$ lautet, wobei η_j die Komponenten von \mathbf{y} bzgl. der Orthonormalbasis $\mathbf{v}_j, j = 1, 2, 3$ sein sollen. Das sieht man so ein:

$$\mathbf{x} \cdot \mathbf{y} = (\xi_1\mathbf{v}_1 + \xi_2\mathbf{v}_2 + \xi_3\mathbf{v}_3) \cdot (\eta_1\mathbf{v}_1 + \eta_2\mathbf{v}_2 + \eta_3\mathbf{v}_3).$$

Jetzt „multipliziert man wieder aus“ und nutzt die Orthogonalität sowie $\mathbf{v}_j \cdot \mathbf{v}_j = 1$ aus. ■

Eine Übertragung dieses Vorgehens auf den \mathbb{R}^n (z.B. auf die Ebene \mathbb{R}^2) ist offensichtlich und wird daher nicht durchgeführt.

2.4 Geraden: Hessesche Normalform

2.4.1 Lot auf Geraden

Gegeben seien eine Gerade $\mathcal{G} = \{P + \lambda \mathbf{v} : \lambda \in \mathbb{R}\}$ mit dem Punkt $P \in \mathcal{G}$ und Richtungsvektor \mathbf{v} , s. Kap. 1.7. Nun sei ein Punkt $Q \notin \mathcal{G}$ gegeben. Wodurch ist der Fußpunkt L des Lotes auf \mathcal{G} bestimmt? Was ist überhaupt ein Lot?

Zwei Dinge erinnern Sie vielleicht aus der Schulzeit:

- Das Lot, d.h. der Verbindungsvektor \vec{LQ} , steht senkrecht auf \mathcal{G} , was in unserer Sprache heißt, dass

$$\vec{LQ} \perp \mathbf{v}.$$

- Unter allen Punkten $P' \in \mathcal{G}$ ist L derjenige Punkt, der zu Q den kürzesten Abstand hat.

Die Gleichwertigkeit dieser beiden Sichtweisen kann man beweisen — was hier nicht geschehen soll. Stattdessen werden wir eine „Formel“ für L herleiten, indem $\vec{LQ} \perp \mathbf{v}$ benutzt wird: Da $L \in \mathcal{G}$ ist, gibt es ein $\lambda \in \mathbb{R}$ mit $L = P + \lambda \mathbf{v}$. Dann ist¹¹ $\vec{LQ} = Q - L = Q - P - \lambda \mathbf{v}$ und $\vec{LQ} \perp \mathbf{v}$ liefert

$$(Q - P - \lambda \mathbf{v}) \cdot \mathbf{v} = 0,$$

woraus $\lambda = \frac{(Q-P) \cdot \mathbf{v}}{|\mathbf{v}|^2}$ und die gesuchte Formel

$$L = P + \frac{(Q - P) \cdot \mathbf{v}}{|\mathbf{v}|^2} \mathbf{v}$$

wird.

Interessieren wir uns für den Richtungsvektor \vec{LQ} , so erhalten wir

$$\vec{LQ} = Q - P - \frac{(Q - P) \cdot \mathbf{v}}{|\mathbf{v}|^2} \mathbf{v}$$

als einen Vektor, der orthogonal zu \mathbf{v} ist. Ist \mathcal{G} eine Gerade in der Ebene \mathbb{R}^2 , so sollte dieser Vektor bis auf Länge und Orientierung eindeutig, also unabhängig von Q sein, man erhält einen sogenannten *Normalenvektor*.

2.4.2 Geradendarstellung mit Hilfe des Skalarprodukts; Hessesche Normalform

In Kap. 1.7.1 hatten wir die *Koordinatenform* $\mathcal{G} := \{(x, y) \in \mathbb{R}^2 : ax + by = c\}$ von in \mathbb{R}^2 gelegenen Geraden kennen gelernt. Mit dem Vektor $\mathbf{a} := (a, b)$, einem zu einem Punkt $X \in \mathcal{G}$

¹¹Eigentlich sollte statt P der Ortsvektor \mathbf{P} und statt Q der Ortsvektor \mathbf{Q} in den folgenden Formeln auftreten. Aber wir wollen ab jetzt etwas großzügiger sein und nicht mehr systematisch zwischen Punkten und ihren zugehörigen Ortsvektoren unterscheiden.

gehörenden Ortsvektor \mathbf{x} und dem Skalarprodukt liest sich das so: $\mathcal{G} := \{X \in \mathbb{R}^2 : \mathbf{x} \cdot \mathbf{a} = c\}$. Nun vergleichen wir diese Darstellung mit der Parameterdarstellung $\mathcal{G} = \{P + \lambda \mathbf{v} : \lambda \in \mathbb{R}\}$. Es kann $X = P + \lambda \mathbf{v}$ gesetzt werden, so dass $(P + \lambda \mathbf{v}) \cdot \mathbf{a} = c$ für alle $\lambda \in \mathbb{R}$ gilt! Das ist aber nur möglich, wenn $\mathbf{v} \cdot \mathbf{a} = 0$, d.h., wenn $\mathbf{a} \perp \mathbf{v}$, d.h., wenn \mathbf{a} *Normalenvektor* der Geraden \mathcal{G} ist.

Beispiel: Sei

$$\mathcal{G} = \{(x, y) \in \mathbb{R}^2 : 3x - 4y = 2\} \quad (2.4)$$

Dann wissen wir sofort, dass der Vektor $\mathbf{a} = (3, -4)$ orthogonal zum Richtungsvektor von \mathcal{G} ist!

Der Vektor $\mathbf{a} = (a, b)$ in $\mathcal{G} := \{(x, y) \in \mathbb{R}^2 : ax + by = c\}$ ist nicht eindeutig durch die Gerade \mathcal{G} bestimmt, da die Gleichung $ax + by = c$ mit irgendwelcher Zahl $\neq 0$ multipliziert werden kann, ohne dass sich etwas an der Punktmenge \mathcal{G} ändert. Warum also nicht \mathbf{a} so normieren, dass $|\mathbf{a}| = 1$? In unserem Beispiel (2.4) können wir auch statt „ $3x - 4y = 2$ “ auch „ $\frac{3}{5}x - \frac{4}{5}y = \frac{2}{5}$ “ schreiben. Dann ist $\mathbf{a} = (\frac{3}{5}, -\frac{4}{5})$ ein normierter Vektor der Länge Eins¹².

Definition 2.10. Die Koordinatenform einer Geraden $\mathcal{G} := \{(x, y) \in \mathbb{R}^2 : ax + by = c\}$ heißt **Hessesche Normalform**, wenn $|\mathbf{a}| = 1$ für $\mathbf{a} = (a, b)$.

Der Vorteil der Hesseschen Normalform wird in folgendem Satz wiedergegeben:

Satz 2.11. Sei $\mathcal{G} := \{(x, y) \in \mathbb{R}^2 : ax + by = c\}$ eine Hessesche Normalform der Geraden \mathcal{G} , d.h. es sei $|\mathbf{a}| = 1$ für $\mathbf{a} := (a, b)$. Dann ist \mathbf{a} ein Normalenvektor der Länge Eins und $|c|$ ist der Abstand der Geraden zum Ursprung.

Letzters müssen wir noch zeigen: Wenn $c = 0$, so ist der Ursprung ein Punkt auf der Geraden, dessen Abstand zur Geraden natürlich Null ist. In diesem Fall ist die Behauptung richtig. Sei also $c \neq 0$. Der Fußpunkt L des Lotes, das vom Ursprung auf \mathcal{G} gefällt wird, hat als Richtungsvektor den Normalenvektor \mathbf{a} , d.h. es muss mit einem $\alpha \in \mathbb{R}$ gelten, dass $L = (\ell_1, \ell_2) = \alpha(a, b)$. Da $L \in \mathcal{G}$, gilt $\ell_1 a + \ell_2 b = c$, also $\alpha(a^2 + b^2) = c$. Wegen $|\mathbf{a}| = 1$ gilt $a^2 + b^2 = 1$ und demnach $\alpha = c$. Der gesuchte Abstand ist $|\mathbf{L}|$, wenn \mathbf{L} der zu L gehörende Ortsvektor ist. Nun ist aber $|\mathbf{L}| = |\alpha| |\mathbf{a}| = |\alpha| = |c|$. ■

2.4.3 Umrechnungen von Koordinaten- in Parameterform und umgekehrt

Wir haben gesehen, dass Richtungsvektor \mathbf{v} und Koeffizientenvektor \mathbf{a} einer Koordinatenform der Beziehung $\mathbf{a} \perp \mathbf{v}$ genügen. Im \mathbb{R}^2 haben wir einen einfachen

Satz 2.12. Zu $\mathbf{v} := (v_1, v_2) \in \mathbb{R}^2$ sind $\mathbf{a} := (-v_2, v_1)$ oder $\mathbf{a} := (v_2, -v_1)$ die einzigen Vektoren, die orthogonal zu \mathbf{v} sind ($\mathbf{v} \perp \mathbf{a}$) und gleiche Länge wie \mathbf{v} haben ($|\mathbf{a}| = |\mathbf{v}|$).

¹²I.A. führt die Normierung zu „hässlichen“ Zahlen.

Ist eine Parameterform (in Kurzform $\mathbf{x} = P + \lambda\mathbf{v}$) gegeben, so kann man mit dieser Formel sofort einen Koeffizientenvektor \mathbf{a} einer Koordinatenform berechnen und $c := P \cdot \mathbf{a}$ setzen.

Ist eine Koordinatenform (in Kurzform $\mathbf{a} \cdot \mathbf{x} = c$) gegeben, so kann man mit derselben Formel einen Richtungsvektor \mathbf{v} berechnen und irgendeinen Punkt P ermitteln, der $P \cdot \mathbf{a} = c$ genügt.

2.5 Ebenen: Hessesche Normalform

Wie bei Geraden könnten wir uns erst einmal mit Loten durch einen Punkt auf eine Ebene befassen. Das tun wir nicht, sondern steuern sofort die Koordinatenform

$$\mathcal{E} := \{(x, y, z) \in \mathbb{R}^3 : ax + by + cz = d\}$$

einer Ebene \mathcal{E} im \mathbb{R}^3 aus Kap. 1.8.5 an. Mit dem Vektor $\mathbf{a} := (a, b, c)$, einem zu einem Punkt $X \in \mathcal{E}$ gehörenden Ortsvektor \mathbf{x} und dem Skalarprodukt liest sich das so:

$$\mathcal{E} := \{X \in \mathbb{R}^3 : \mathbf{x} \cdot \mathbf{a} = d\}.$$

Wir werden jetzt zeigen, dass \mathbf{a} orthogonal zu jedem Richtungsvektor \mathbf{v}, \mathbf{w} von \mathcal{E} ist. Um dies zu zeigen, gehe ich von einer Parameterdarstellung

$$\mathcal{E} = \{P + \lambda\mathbf{v} + \mu\mathbf{w} : \lambda, \mu \in \mathbb{R}\}$$

von \mathcal{E} aus. Mit $X = P + \lambda\mathbf{v} + \mu\mathbf{w}$ ergibt sich

$$(P + \lambda\mathbf{v} + \mu\mathbf{w}) \cdot \mathbf{a} = d \quad \forall \lambda, \mu \in \mathbb{R}.$$

Da $P \in \mathcal{E}$, gilt $P \cdot \mathbf{a} = d$ und unter Ausnutzung der Rechenregeln des Skalarprodukts folgt

$$(\lambda\mathbf{v} + \mu\mathbf{w}) \cdot \mathbf{a} = 0 \quad \forall \lambda, \mu \in \mathbb{R}.$$

Das kann nur sein, wenn sowohl $\mathbf{v} \perp \mathbf{a}$ als auch $\mathbf{w} \perp \mathbf{a}$.

Wieder können wir \mathbf{a} normieren und erhalten analog zu Satz 2.11

Satz 2.13. *Sei $\mathcal{E} := \{(x, y, z) \in \mathbb{R}^3 : ax + by + cz = d\}$ eine Hessesche Normalform, d.h. es sei $|\mathbf{a}| = 1$ für $\mathbf{a} := (a, b, c)$. Dann ist \mathbf{a} ein Normalenvektor der Länge Eins und $|d|$ ist der Abstand der Ebene zum Ursprung.*

Der Beweis verläuft ganz ähnlich wie der zu Satz 2.11.

Wie macht man aber aus irgendeiner Koordinatenform $ax + by + cz = d$ eine Hessesche Normalform? Ganz einfach: Man „normiert“ den Koeffizientenvektor $\mathbf{a} = (a, b, c)$ auf Länge Eins, indem man die Gleichung $ax + by + cz = d$ durch die Länge $|\mathbf{a}|$ dividiert. Erinnern Sie sich: Wenn $\mathbf{a} \neq \mathbf{o}$, so hat $\frac{\mathbf{a}}{|\mathbf{a}|}$ die Länge Eins.

Beispiel: Aus einer Koeffizientenform $x - 2y + z = 1$ wird wegen $\mathbf{a} = (1, -2, 1)$ und $|\mathbf{a}| = \sqrt{6}$ eine Hessesche Normalform

$$\frac{1}{\sqrt{6}}x - \frac{2}{\sqrt{6}}y + \frac{1}{\sqrt{6}}z = \frac{1}{\sqrt{6}},$$

und damit ist $\frac{1}{\sqrt{6}}$ der Abstand der Ebene zum Ursprung.

2.5.1 Umrechnungen von Koordinaten- in Parameterform und umgekehrt

ZU einer gegebenen Koordinatenform (Kurzform $\mathbf{a} \cdot \mathbf{x} = d$) eine Parameterform zu berechnen, ist relativ einfach. Man bestimmt irgendeinen Punkt $P \in \mathbb{R}^3$, der $P \cdot \mathbf{a} = d$ genügt und zwei linear unabhängige \mathbf{v} und \mathbf{w} , die orthogonal zu \mathbf{a} sind. Dabei kann man den Trick von Satz 2.12 verwenden, wenn man in \mathbf{v} und \mathbf{w} jeweils eine Komponente Null setzt.

Beispiel: Ist $\mathbf{a} := (1, -2, 3)$, so sind $\mathbf{v} := (0, 3, 2)$ und $\mathbf{w} := (-3, 0, 1)$ hierzu orthogonal.

Schwieriger erscheint schon die Umrechnung von Parameterform in Koordinatenform. Offensichtlich ist zu linear unabhängigen \mathbf{v} und \mathbf{w} ein dritter Vektor $\mathbf{a} \neq \mathbf{o}$ gesucht, der auf \mathbf{v} und \mathbf{w} senkrecht steht. Das nächste Kapitel gibt die Antwort: Man setze

$$\mathbf{a} := \mathbf{v} \times \mathbf{w} = (v_2w_3 - v_3w_2, v_3w_1 - v_1w_3, v_1w_2 - v_2w_1).$$

Die rechte Seite d der Koordinatenkurzform $\mathbf{a} \cdot \mathbf{x} = d$ berechnet sich einfach aus $d := P \cdot \mathbf{a}$.

2.6 Vektorprodukt

Das im Folgenden behandelte *Vektorprodukt* hat nur im dreidimensionalen euklidischen Raum \mathbb{R}^3 Bedeutung. Es spielt eine große Rolle in der Physik, hat aber auch Anwendungen in der (Computer-) Geometrie. So kann man aus der Koordinatenform einer Ebene mittels Vektorprodukt aus einem Richtungspaar einen Normalenvektor der Ebene berechnen!

Es kann sein, dass ich dieses Kapitel in der Vorlesung nicht im gesamten Umfang behandelte.

Die Aufgabe, die wir uns jetzt stellen, lautet: Gegeben seien zwei Vektoren \mathbf{a} und \mathbf{b} des \mathbb{R}^3 , beide $\neq \mathbf{o}$. Gesucht ist ein dritter Vektor \mathbf{c} , den man mit Hilfe einer noch zu definierenden binären Verknüpfung $\mathbf{c} = \mathbf{a} \times \mathbf{b}$ (sprich „ \mathbf{a} kreuz \mathbf{b} “ oder **Vektorprodukt** aus \mathbf{a} und \mathbf{b}) berechnet, mit den Eigenschaften

- $(\mathbf{a} \times \mathbf{b}) \perp \mathbf{a}, \quad (\mathbf{a} \times \mathbf{b}) \perp \mathbf{b}$
- $|\mathbf{a} \times \mathbf{b}| = A$, wobei A die Fläche des von \mathbf{a} und \mathbf{b} aufgespannten Parallelogramms ist, siehe Kap. 2.2.3.

Wegen Satz 2.6 gilt also

$$|\mathbf{a} \times \mathbf{b}| = |\mathbf{a}||\mathbf{b}| \sin(\angle(\mathbf{a}, \mathbf{b})).$$

Geometrisch ist klar, dass \mathbf{c} hier bis auf das Vorzeichen eindeutig festgelegt ist. Es wird die „Orientierung“ von \mathbf{c} so festgelegt, dass die „Rechte-Hand-Regel“ in Abb. 2.10 gilt:

Sind Mittelfinger entlang \mathbf{a} und Daumen entlang \mathbf{b} orientiert, so stellt der Zeigefinger die Orientierung des Vektorprodukts $\mathbf{a} \times \mathbf{b}$ dar. Wird \mathbf{a} auf kürzestem Weg in \mathbf{b} hineingedreht, so

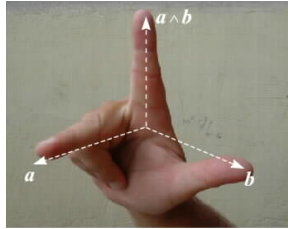


Abbildung 2.10: Rechte-Hand-Regel

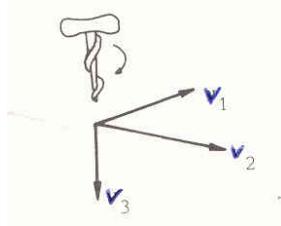


Abbildung 2.11: Korkenzieher-Regel

entspricht das einer „Rechtsschrauben-Drehung“. $\mathbf{c} = \mathbf{a} \times \mathbf{b}$ zeigt daher in die Bewegungsrichtung der hypothetischen Schraube¹³.

Auch wenn wir noch keine „Formel“ für $\mathbf{a} \times \mathbf{b}$ kennen¹⁴, folgen aus den obigen Forderungen die folgenden Rechenregeln

- $\mathbf{a} \times \mathbf{a} = \mathbf{o}$, da ein Parallelogramm, dessen beide Seiten zusammenfallen, die Fläche Null hat.
- **Schiefsymmetrie:** $\mathbf{a} \times \mathbf{b} = -\mathbf{b} \times \mathbf{a}$. Dies folgt aus der „Rechte-Hand-Regel“.
- Ist $\lambda \in \mathbb{R}$ ein Skalar, so gilt $(\lambda \cdot \mathbf{a}) \times \mathbf{b} = \mathbf{a} \times (\lambda \cdot \mathbf{b}) = \lambda \cdot (\mathbf{a} \times \mathbf{b})$. Dies folgt aus der Tatsache, dass $|\Delta(\mathbf{a}, \mathbf{b})|$ die Fläche des von \mathbf{a} und \mathbf{b} aufgespannten Parallelogramms ist und dass $\Delta(\lambda \cdot \mathbf{a}, \mathbf{b}) = \lambda \Delta(\mathbf{a}, \mathbf{b})$, was zur offensichtlichen Konsequenz führt, dass sich die Fläche gemäß λ vervielfacht, wenn sich eine Kante „ver- λ -facht“.
- $\mathbf{a} \times (\mathbf{b} + \mathbf{c}) = \mathbf{a} \times \mathbf{b} + \mathbf{a} \times \mathbf{c}$ — eine Art Distributivität. Der Beweis ist nicht offensichtlich. Als erstes muss man zeigen, dass $\mathbf{a} \times \mathbf{b} + \mathbf{a} \times \mathbf{c}$ sowohl auf \mathbf{a} als auch auf $\mathbf{b} + \mathbf{c}$ senkrecht steht. Analog
- $(\mathbf{a} + \mathbf{b}) \times \mathbf{c} = \mathbf{a} \times \mathbf{c} + \mathbf{b} \times \mathbf{c}$
- $\mathbf{a} \times \mathbf{o} = \mathbf{o}$

¹³Denken Sie an einen Korkenzieher! S. Abb. 2.11

¹⁴Diese finden Sie in (2.6.1).

- \mathbf{a} ist parallel zu \mathbf{b} genau dann, wenn $\mathbf{a} \times \mathbf{b} = \mathbf{o}$. Die letzten beiden Eigenschaften ergeben sich wieder aus den Flächenbetrachtungen des von \mathbf{a} und \mathbf{b} aufgespannten Parallelogramms.

2.6.1 Berechnung des Vektorprodukts

Nun wollen wir eine Formel für $\mathbf{a} \times \mathbf{b}$ mit $\mathbf{a} = (a_1, a_2, a_3)$ und $\mathbf{b} = (b_1, b_2, b_3)$ gewinnen. Dazu nehmen wir an, dass $\mathbf{e}_1 \times \mathbf{e}_2 = \mathbf{e}_3$, $\mathbf{e}_1 \times \mathbf{e}_3 = -\mathbf{e}_2$ und $\mathbf{e}_2 \times \mathbf{e}_3 = \mathbf{e}_1$.

Diese Vereinbarung ist für den nachfolgenden Satz 2.15 von Belang.

Nun gilt

$$\mathbf{a} = a_1\mathbf{e}_1 + a_2\mathbf{e}_2 + a_3\mathbf{e}_3, \quad \mathbf{b} = b_1\mathbf{e}_1 + b_2\mathbf{e}_2 + b_3\mathbf{e}_3.$$

Aus den Distributivgesetzen des Vektorprodukts und $\mathbf{e}_1 \times \mathbf{e}_2 = \mathbf{e}_3$, $\mathbf{e}_1 \times \mathbf{e}_3 = -\mathbf{e}_2$, $\mathbf{e}_2 \times \mathbf{e}_3 = \mathbf{e}_1$ folgt

$$\begin{aligned} \mathbf{a} \times \mathbf{b} &= (a_2b_3 - a_3b_2)\mathbf{e}_1 + (a_3b_1 - a_1b_3)\mathbf{e}_2 + \\ &\quad (a_1b_2 - a_2b_1)\mathbf{e}_3. \end{aligned}$$

Also in Koordinatenschreibweise

$$\mathbf{a} \times \mathbf{b} = (a_2b_3 - a_3b_2, a_3b_1 - a_1b_3, a_1b_2 - a_2b_1). \quad (2.5)$$

Beispiel: $\mathbf{a} = (2, 4, 1)$, $\mathbf{b} = (3, -2, 1)$. Dann ist $\mathbf{c} := \mathbf{a} \times \mathbf{b} = (6, 1, -16)$. Überzeugen Sie sich, dass $\mathbf{c} \perp \mathbf{a}$ und $\mathbf{c} \perp \mathbf{b}$. Die Formel

$$A^2 = |\mathbf{a}|^2|\mathbf{b}|^2 - (\mathbf{a} \cdot \mathbf{b})^2$$

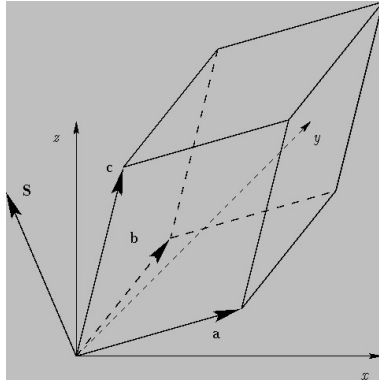
für die Fläche A des von \mathbf{a} und \mathbf{b} aufgespannten Parallelogramms liefert $A^2 = 21 \cdot 14 - 1 = 293$ in Übereinstimmung mit $|\mathbf{c}|^2 = 36 + 1 + 256 = 293$.

Man kann jetzt leicht nachrechnen, dass das Vektorprodukt $\mathbf{a} \times \mathbf{b}$ alle genannten Eigenschaften hat! Insbesondere, dass $\mathbf{a} \times \mathbf{b} \perp \mathbf{a}$ und $\mathbf{a} \times \mathbf{b} \perp \mathbf{b}$. Nur die Orientierung („Rechte-Hand-Regel“) bedarf einer Begründung. Diese ist gegeben, wenn das zugrundeliegende Koordinatensystem ein „Rechtssystem“ ist. Dieses bedeutet, dass die Richtung der z -Achse diejenige Richtung ist, in die ein Korkenzieher hineindreht, wenn er die x -Achse um 90 Grad in die y -Achse hineindreht.

Aufgabe: Überprüfen Sie, welche der Koordinatensysteme in Abb. 1.7-1.16 Rechts-Systeme sind. Das in der Vorlesung verwendete System ist ein Linkssystem!

Bemerkung: Das in der Vorlesung verwendete Koordinatensystem ist ein Linkssystem!

Jetzt kann man ebenfalls leicht nachrechnen, dass das Vektorprodukt keine assoziative Verknüpfung ist. Z.B. gilt mit $\mathbf{c} = \mathbf{e}_1 = (1, 0, 0)$ und $\mathbf{a} = (2, 4, 1)$, $\mathbf{b} = (3, -2, 1)$ aus dem Beispiel eben, dass $(\mathbf{a} \times \mathbf{b}) \times \mathbf{c} = (6, 1, -16) \times (1, 0, 0) = (0, -16, -1)$ und $\mathbf{a} \times (\mathbf{b} \times \mathbf{c}) = (2, 4, 1) \times (0, 1, 2) = (7, -4, 2)$.

Abbildung 2.12: Parallelepiped: $\mathbf{S} = \mathbf{a} \times \mathbf{b}$

2.6.2 Vektorprodukt und Determinanten von 2x2-Matrizen

Wir wissen aus Kap. 1.5.3, dass

$$\Delta(\mathbf{a}, \mathbf{b}) = a_1 b_2 - a_2 b_1$$

die Determinante der 2×2 -Matrix

$$\begin{pmatrix} a_1 & b_1 \\ a_2 & b_2 \end{pmatrix}$$

ist. Die Berechnungsformel (2.5) für das Vektorprodukt $\mathbf{a} \times \mathbf{b}$ lässt erkennen, dass jede Komponente des Vektorprodukts die Gestalt einer solchen Determinante hat. Wir führen die Schreibweise

$$\Delta(\mathbf{a}, \mathbf{b}) := \begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix}$$

ein und können damit schreiben

$$\mathbf{a} \times \mathbf{b} = \begin{vmatrix} a_2 & b_2 \\ a_3 & b_3 \end{vmatrix} \mathbf{e}_1 - \begin{vmatrix} a_1 & b_1 \\ a_3 & b_3 \end{vmatrix} \mathbf{e}_2 + \begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix} \mathbf{e}_3$$

Beispiel: $\mathbf{e}_1 \times \mathbf{b} = (0, -b_3, b_2)$.

2.6.3 Das Volumen eines Parallelepipedes

Die dreidimensionale Verallgemeinerung eines Parallelogramms ist ein von drei Kantenvektoren \mathbf{a} , \mathbf{b} und \mathbf{c} aufgespanntes **Parallelepiped**, s. Abb. 2.12.

So, wie ein Rechteck ein Spezialfall eines Parallelogramms ist, so ist ein *Quader* (mit drei paarweise aufeinander senkrecht stehenden Vektoren) ein Spezialfall eines Parallelepipedes. Das Volumen V eines Parallelepipedes ist durch die Formel „Grundfläche x Höhe“ gegeben. Die von \mathbf{a} und \mathbf{b} aufgespannte Grundfläche ist ja gerade nach Konstruktion $|\mathbf{a} \times \mathbf{b}|$, der Höhenvektor steht senkrecht auf \mathbf{a} und \mathbf{b} . Die Höhe h des Parallelepipedes errechnet sich zu $h = |\mathbf{c}| |\cos(\varphi)|$,

wenn φ der Winkel zwischen \mathbf{c} und einer Normalen zur Ebene des aus \mathbf{a} und \mathbf{b} bestehenden Parallelogramms ist. Damit ergibt sich

$$V = h|\mathbf{a} \times \mathbf{b}|.$$

Nach Konstruktion ist $\mathbf{a} \times \mathbf{b}$ orthogonal zu \mathbf{a} und \mathbf{b} , und es ist

$$|(\mathbf{a} \times \mathbf{b}) \cdot \mathbf{c}| = |\mathbf{a} \times \mathbf{b}||\mathbf{c}|\cos(\varphi) = V.$$

bzw.

$$V = |(\mathbf{a} \times \mathbf{b}) \cdot \mathbf{c}|.$$

Wie erwartet, ist $V = 0$, wenn das aus \mathbf{a} , \mathbf{b} und \mathbf{c} bestehende System linear abhängig ist. Z.B. dann, wenn \mathbf{c} in der von \mathbf{a} und \mathbf{b} aufgespannten Ebene liegt, da dann $(\mathbf{a} \times \mathbf{b}) \perp \mathbf{c}$. Jetzt haben wir eine einfache Methode, die lineare Unabhängigkeit von drei Vektoren im \mathbb{R}^3 zu prüfen:

Satz 2.14. *Das aus \mathbf{a} , \mathbf{b} und \mathbf{c} bestehende System ist linear abhängig genau dann, wenn*

$$(\mathbf{a} \times \mathbf{b}) \cdot \mathbf{c} = 0.$$

Man kann jetzt auch ohne „Rechte-Hand-Regel“ entscheiden, ob ein aus drei Vektoren \mathbf{a} , \mathbf{b} und \mathbf{c} bestehendes System ein Rechts- oder Links-System ist (vorausgesetzt, das zugrundeliegende Koordinatensystem ist ein Rechtssystem):

Satz 2.15. *Ein aus drei Vektoren \mathbf{a} , \mathbf{b} und \mathbf{c} bestehendes System ist ein Rechts- (bzw. ein Links-) System genau dann, wenn $(\mathbf{a} \times \mathbf{b}) \cdot \mathbf{c} > 0$ bzw. < 0 .*

Man kann also auch für das aus drei Vektoren \mathbf{a} , \mathbf{b} und \mathbf{c} bestehende linear unabhängige System von einem *Rechtssystem* reden, wenn der Winkel zwischen \mathbf{c} und $\mathbf{a} \times \mathbf{b}$ kleiner als ein rechter Winkel ist.

Man nennt $(\mathbf{a} \times \mathbf{b}) \cdot \mathbf{c}$ auch das **orientierte Volumen** eines Parallelepipeds oder das **Spatprodukt** der drei Vektoren.

Übrigens kann man jetzt auch die Formel für die Fläche eines Parallelogramms aus Satz 1.13, $A = |a_1b_2 - a_2b_1|$, wenn $\mathbf{a} = (a_1, a_2)$, $\mathbf{b} = (b_1, b_2)$ „wiederentdecken“. Dazu nehme man die dreidimensionale Version $\mathbf{a} = (a_1, a_2, 0)$ und $\mathbf{b} = (b_1, b_2, 0)$ der beiden Vektoren und betrachte das Parallelepipeds mit \mathbf{a} , \mathbf{b} und $\mathbf{c} := \mathbf{e}_3$. Dessen Volumen ist ja gerade gleich der Fläche des Parallelogramms. Nun ist $(a_1, a_2, 0) \times (b_1, b_2, 0) = (0, 0, a_1b_2 - a_2b_1)$ und $(\mathbf{a} \times \mathbf{b}) \cdot \mathbf{c} = a_1b_2 - a_2b_1$.

2.6.4 Determinante einer 3x3-Matrix

Die drei Vektoren eines Parallelepipeds kann man zu einer 3×3 -Matrix zusammenstellen — mit \mathbf{a} in der ersten, \mathbf{b} in der zweiten und \mathbf{c} in der dritten Spalte:

$$M := \begin{pmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{pmatrix}$$

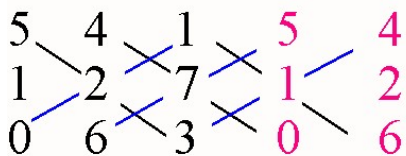


Abbildung 2.13: Regel von Sarrus

Definition 2.16. $\Delta(M) = \Delta(\mathbf{a}, \mathbf{b}, \mathbf{c}) = (\mathbf{a} \times \mathbf{b}) \cdot \mathbf{c}$ heißt die **Determinante** von M .

Nun liest sich Satz 2.14 auch so:

Satz 2.17. Das aus \mathbf{a} , \mathbf{b} und \mathbf{c} bestehende System ist linear unabhängig genau dann, wenn die Determinante $\Delta(\mathbf{a}, \mathbf{b}, \mathbf{c}) \neq 0$, bzw. wenn die Determinante der aus den drei Vektoren gebildete Matrix M nicht verschwindet.

Der Vollständigkeit halber geben wir die Berechnungsvorschrift an:

$$\Delta(\mathbf{a}, \mathbf{b}, \mathbf{c}) = a_1 b_2 c_3 + a_2 b_3 c_1 + a_3 b_1 c_2 - a_1 b_3 c_2 - a_2 b_1 c_3 - a_3 b_2 c_1.$$

Sie spüren wahrscheinlich eine gewisse Systematik dieser Berechnungsformel. Diese wird deutlich durch die sog. *Jägerzaunregel* oder *Regel von Sarrus*, die ich hier aber nicht angebe, sondern auf [Erklärung der Regel von Sarrus](#) (mathematik.net) verweise.

Als Beispiel schauen wir uns Abb. 2.13 an:

Die ersten beiden Spalten werden (in Rot) an die Matrix angehängt. Dann werden drei „Hauptdiagonalprodukte“ (in schwarz) berechnet und addiert ($5 \cdot 2 \cdot 3 + 4 \cdot 7 \cdot 0 + 1 \cdot 1 \cdot 6$) und von diesen die Summe der drei „Nebendiagonalprodukte“ (in blau) subtrahiert. Wenn Sie richtig rechnen, kommt $\Delta(\mathbf{a}, \mathbf{b}, \mathbf{c}) = -186$ heraus.

2.6.5 Internet

Unter [Applet Vektorprodukt](#) (Moltke-Homepage) kann man das Vektorprodukt veranschaulichen.

Noch schöner ist das englischsprachige Applet [Vector Cross Product](#) (Syracuse UUniversity USA) Hier kann man mit der Maus das Koordinatensystem drehen und die Lage und Länge der Vektoren ändern.

Kapitel 3

Lineare Abbildungen

3.1 Einführung

In diesem Kapitel betrachten wir *lineare Abbildungen* $\mathcal{A} : \mathbb{R}^n \rightarrow \mathbb{R}^m$, wobei *Linearität* durch die Eigenschaft

$$\mathcal{A}(\mathbf{u} + \mathbf{v}) = \mathcal{A}(\mathbf{u}) + \mathcal{A}(\mathbf{v}) \quad (3.1)$$

$$\mathcal{A}(\lambda \mathbf{u}) = \lambda \mathcal{A}(\mathbf{u}) \quad (3.2)$$

für alle $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$ und für alle Skalare $\lambda \in \mathbb{R}$ ausgedrückt wird. Es wird sich zeigen, dass sich mit solch einer Abbildung stets eine $m \times n$ -Matrix A verbindet¹, das ist eine rechteckige Tabelle mit m Zeilen und n Spalten,

$$A := \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}, \quad (3.3)$$

deren j -te Spalten(-Vektoren) \mathbf{a}_j übrigens gerade $\mathcal{A}(\mathbf{e}_j)$, $j = 1, 2, \dots, n$ sind. Solche Matrizen treten ganz natürlich bei *linearen Gleichungssystemen (LGS)*

$$a_{j1}x_1 + a_{j2}x_2 + \cdots + a_{jn}x_n = b_j, \quad j = 1, 2, \dots, m \quad (3.4)$$

auf, deren Lösung $\mathbf{x} \in \mathbb{R}^n$ die Gleichung $\mathcal{A}(\mathbf{x}) = \mathbf{b}$ mit einem gegebenen Vektor $\mathbf{b} \in \mathbb{R}^m$ erfüllt — gesucht ist also ein Urbild von \mathbf{b} unter der Abbildung \mathcal{A} . Hierbei handelt es sich um m Gleichungen für n Unbekannte x_k , $k = 1, 2, \dots, n$.

Eine Lösungstheorie (Existenz, Eindeutigkeit) zu diesen Gleichungen (der wichtigste Fall ist $m = n$) wird zumindestens angedeutet.

¹Wir schreiben auch $(m \times n)$ -Matrix

Ein anderer Zugang zu linearen Abbildungen ist der als *geometrische Transformationen* wie es sie z.B. in der Computergrafik gibt. Hier werden wir einfache Rotationen, Spiegelungen und Scherungen in der Ebene beschreiben.

Die bei der Zentralprojektion in der darstellenden Geometrie sich ergebende Abbildung ist nicht linear. Sie ist in den Übungen behandelt worden, und soll hier auch kurz erwähnt werden.

An Webseiten erwähne ich:

- [Leitprogramm Lineare Gleichungssysteme](#) (ETH Zürich)

3.2 Lineare Gleichungssysteme

3.2.1 Bisherige Beispiele

In Kap. 1.5.3 haben wir zu gegebenem $\mathbf{a} = (a_1, a_2) \in \mathbb{R}^2$ (und zu $\mathbf{v}_1 = (v_{11}, v_{21}), \mathbf{v}_2 = (v_{12}, v_{22}) \in \mathbb{R}^2$) die beiden Linearkoeffizienten x_1 und x_2 der Linearkombination von $\mathbf{a} = x_1\mathbf{v}_1 + x_2\mathbf{v}_2$ durch die $\mathbf{v}_j, j = 1, 2$, ausgeschrieben als zwei lineare Gleichungen

$$v_{11}x_1 + v_{12}x_2 = a_1, \quad v_{21}x_1 + v_{22}x_2 = a_2, \quad (3.5)$$

gesucht — dies ist ein **lineares Gleichungssystem** mit 2 Gleichungen für die beiden Unbekannten x_1, x_2 . Wir haben die Tabelle

$$V := \begin{pmatrix} v_{11} & v_{12} \\ v_{21} & v_{22} \end{pmatrix}$$

die **Koeffizientenmatrix** von (3.5) und den Ausdruck

$$\Delta(V) := v_{11}v_{22} - v_{12}v_{21} \quad (3.6)$$

die **Determinante** von V genannt. Es ergab sich der folgende

Satz 3.1. *Die folgenden Aussagen sind gleichwertig:*

- (L1) *Das System der beiden Vektoren \mathbf{v}_1 und \mathbf{v}_2 ist linear unabhängig.*
- (L2) *Die Determinante $\Delta(V)$ ist ungleich Null.*
- (L3) *Das lineare Gleichungssystem (3.5) hat eine eindeutige Lösung.*

In Kap. 1.7.4 wurde auch skizziert, wie man mit Hilfe eines Eliminationsverfahrens dieses lineare Gleichungssystem löst.

In Kap. 2.6.4 haben wir die drei Vektoren eines Parallelepipedes zu einer 3×3 -Matrix zusammengestellt — mit \mathbf{a} in der ersten, \mathbf{b} in der zweiten und \mathbf{c} in der dritten Spalte:

$$M := \begin{pmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{pmatrix}$$

Mit $\Delta(M) = \Delta(\mathbf{a}, \mathbf{b}, \mathbf{c}) = (\mathbf{a} \times \mathbf{b}) \cdot \mathbf{c}$ wurde die **Determinante** von M bezeichnet, für die sich $\Delta(\mathbf{a}, \mathbf{b}, \mathbf{c}) = a_1b_2c_3 + a_2b_3c_1 + a_3b_1c_2 - a_1b_3c_2 - a_2b_1c_3 - a_3b_2c_1$ ergibt. Diese stellt das vorzeichenbehaftete Volumen des Parallelepipeds dar.

Es gilt ein zu Satz 3.1 völlig analoger Satz, wenn wir zu gegebenem $\mathbf{a} = (a_1, a_2, a_3) \in \mathbb{R}^3$ (und zu $\mathbf{v}_1 = (v_{11}, v_{21}, v_{31}), \mathbf{v}_2 = (v_{12}, v_{22}, v_{32}), \mathbf{v}_3 = (v_{13}, v_{23}, v_{33}) \in \mathbb{R}^3$) die drei Linearkoeffizienten x_1, x_2 und x_3 der Linearkombination von $\mathbf{a} = x_1\mathbf{v}_1 + x_2\mathbf{v}_2 + x_3\mathbf{v}_3$ der $\mathbf{v}_j, j = 1, 2, 3$, suchen. Komponentenweise ausgeschrieben ergeben sich drei lineare Gleichungen

$$v_{11}x_1 + v_{12}x_2 + v_{13}x_3 = a_1, \quad v_{21}x_1 + v_{22}x_2 + v_{23}x_3 = a_2, \quad v_{31}x_1 + v_{32}x_2 + v_{33}x_3 = a_3, \quad (3.7)$$

also ein **lineares Gleichungssystem** mit drei Gleichungen und drei Unbekannten. An Stelle von M müssen wir dann die Tabelle

$$V := \begin{pmatrix} v_{11} & v_{12} & v_{13} \\ v_{21} & v_{22} & v_{23} \\ v_{31} & v_{32} & v_{33} \end{pmatrix}$$

betrachtet — eine 3×3 -Matrix. Auch hier liefert die Aussage „Die Determinante von V ist ungleich Null“ ($\Delta(V) \neq 0$) die *Basiseigenschaft* von $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ in \mathbb{R}^3 und damit die eindeutige Lösbarkeit des LGS (3.7).

Allgemein kann man fragen, wann n Vektoren $\mathbf{v}_1, \dots, \mathbf{v}_n \in \mathbb{R}^n$ eine Basis des \mathbb{R}^n liefern. Auch hier wurde in Kap. 1.5 auf die „ $n \times n$ -Matrix“

$$V := \begin{pmatrix} v_{11} & v_{12} & \cdots & v_{1n} \\ v_{21} & v_{22} & \cdots & v_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ v_{n1} & v_{n2} & \cdots & v_{nn} \end{pmatrix}$$

verwiesen.

Benzinherstellung

Die folgende Aufgabe entspricht der Übungsaufgabe 20:

Bei der Rohöl-Bearbeitung entstehen u.a. drei verschiedene Brennstoffe, aus denen Superbenzin mit einer Oktanzahl 100 gemischt werden soll. Die drei Brennstoffe haben die Oktanzahlen 108, 90, 85. Es sollen S Liter Superbenzin hergestellt werden.

Welche linearen „Misch-Gleichungen“ ergeben sich für die verwendeten Mengen x_1, x_2 und x_3 der drei Sorten? Ein Gemisch mit x_1, x_2, x_3 (Litern) hat die Oktanzahl $(x_1 \cdot 108 + x_2 \cdot 90 + x_3 \cdot$

85)/($x_1 + x_2 + x_3$) — dies ist ein gewichteter Mittelwert. Setzt man $S = x_1 + x_2 + x_3$ und fordert man als Ergebnis die Oktanzahl 100, so ergibt sich das lineare Gleichungssystem

$$\begin{aligned} 108x_1 + 90x_2 + 85x_3 &= 100S, \\ x_1 + x_2 + x_3 &= S \end{aligned}$$

mit zwei Gleichungen und drei Unbekannten.

3.2.2 Formulierung eines allgemeinen linearen Gleichungssystems

Die Formulierung eines allgemeinen linearen Gleichungssystems (LGS) in m Gleichungen und n Unbekannten ist jetzt offensichtlich. Wie in vielen Lehrbüchern üblich nennen wir die *Koeffizienten* des LGS a_{jk} , die Unbekannten weiterhin x_k , $k = 1, 2, \dots, n$, und die „rechten Seiten“ b_j , $j = 1, 2, \dots, m$. Die j -te Gleichung lautet damit

$$a_{j1}x_1 + a_{j2}x_2 + \dots + a_{jn}x_n = b_j, \quad j = 1, 2, \dots, m. \quad (3.8)$$

a_{jk} ist also der Koeffizient vor x_k in der j -ten Gleichung. Das LGS heißt *linear*, weil die Unbekannten nur in ihrer „ersten Potenz“ $x_k^1 = x_k$ vorkommen.

Wir stellen die a_{jk} wieder zu einer rechteckigen Tabelle mit m Zeilen und n Spalten zusammen:

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

und nennen A eine $m \times n$ -Matrix. Wir schreiben auch kurz $A := (a_{jk})_{j=1, \dots, m, k=1, \dots, n}$.

Die Menge aller $m \times n$ -Matrizen bezeichnen wir mit $\mathbb{R}^{m \times n}$.

Auf die Matrix A bzw. auf das LGS (3.8) gibt es eine *elementweise Sicht*, eine *Spalten-* und eine *Zeilensicht*. Zur Spaltensicht sehen wir die Spalten von A als Vektoren $\mathbf{a}_k := (a_{1k}, a_{2k}, \dots, a_{mk}) \in \mathbb{R}^m$, $k = 1, 2, \dots, n$. Mit dem Vektor $\mathbf{b} := (b_1, b_2, \dots, b_m) \in \mathbb{R}^m$ kann man (3.8) jetzt als

$$\mathbf{a}_1x_1 + \mathbf{a}_2x_2 + \dots + \mathbf{a}_nx_n = \mathbf{b} \quad (3.9)$$

schreiben und in der uns bekannten Weise so formulieren: Gesucht ist eine Linearkombination des Vektors \mathbf{b} durch Vektoren \mathbf{a}_k , $k = 1, 2, \dots, n$.

Übrigens ist es naheliegend, die Vektoren \mathbf{a}_k , die ja die Spalten der Matrix A bezeichnen, als „Spaltenvektoren“

$$\mathbf{a}_k = \begin{pmatrix} a_{1k} \\ a_{2k} \\ \vdots \\ a_{mk} \end{pmatrix}$$

zu schreiben — wie auch den Vektor

$$\mathbf{b} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}.$$

Das LGS wird dann auch kurz als $A\mathbf{x} = \mathbf{b}$, ausgeschrieben

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix} \quad (3.10)$$

formuliert.

Jetzt kommen wir zur „Zeilensicht“ des LGS (3.8): Hierzu betrachten wir die Zeilen von A als (Zeilen-) Vektoren $\mathbf{a}^j = (a_{j1}, a_{j2}, \dots, a_{jn}) \in \mathbb{R}^n, j = 1, 2, \dots, m$. Dann lautet die j -te Gleichung unter Benutzung des *Skalarprodukts* aus Kap. 2

$$\mathbf{a}^j \cdot \mathbf{x} = b_j, j = 1, 2, \dots, m,$$

wenn man die Unbekannten zu einem Vektor \mathbf{x} zusammenstellt. Hier gibt es gute Gründe, sich \mathbf{x} als Spaltenvektor

$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$

vorzustellen.

Wäre $n = 3$, so wäre

$$\mathcal{E}_j := \{\mathbf{x} \in \mathbb{R}^n : \mathbf{a}^j \cdot \mathbf{x} = b_j\}$$

eine Ebene in \mathbb{R}^3 und die gesuchte Lösung $\mathbf{x} \in \mathbb{R}^3$ könnte man geometrisch als Punkt des Durchschnitts von m solcher Ebenen sehen. Diese geometrische Sichtweise ist sehr anschaulich — liefert es doch interpretierbare Lösungskriterien² — und führt für beliebige n auf die Bezeichnung von \mathcal{E}_j als **Hyperebene** des \mathbb{R}^n . In diesem Sinne kann man das LGS (3.8) als Suche nach gemeinsamen Punkten von m Hyperebenen sehen.

²Drei Ebenen in hinreichend „allgemeiner Lage“ sollten sich in genau einem Punkt schneiden. Mehr als drei Ebenen haben i.A. keinen gemeinsamen Punkt. Zwei Ebenen schneiden sich i.A. in einer Geraden, es sei denn, sie sind parallel.

3.2.3 Anwendungen

Produktionsplan

Wir fangen mit einem betriebswirtschaftlichen Problem an. In einer Firma gäbe es zwei Maschinen M_1 und M_2 , die zur Herstellung von zwei verschiedenen Produkten P_1 und P_2 verwendet werden müssen. Für das Produkt P_k benötigt Maschine M_j eine gewisse Anzahl a_{jk} von Zeiteinheiten (etwa eine Stunde). Es stehen b_j Zeiteinheiten von Maschinen des Typs M_j zur Verfügung. Wieviele Mengeneinheiten x_k des Produkts P_k können alle Maschinen zusammen bei voller Auslastung herstellen?

Die „Lösung“ dieses Problems liegt eigentlich schon in der Aufgabenformulierung und der **Benennung** der gegebenen Daten

$$A := \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$$

und

$$\mathbf{b} := \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}$$

sowie der gesuchten Anzahlen x_1 und x_2 .

Offensichtlich werden für $x_k, k = 1, 2$, Mengeneinheiten (ME) des Produkts P_k gerade $a_{11}x_1 + a_{12}x_2$ Zeiteinheiten Maschinen des Typs M_1 und $a_{21}x_1 + a_{22}x_2$ Zeiteinheiten Maschinen des Typs M_2 benötigt, so dass sich das LGS

$$a_{11}x_1 + a_{12}x_2 = b_1, \quad a_{21}x_1 + a_{22}x_2 = b_2 \quad (3.11)$$

ergibt. Die Lösungskomponenten x_1, x_2 müssen übrigens nicht ganzzahlig sein. Wenn nur Ganzzahligkeit Sinn macht (z.B., wenn es sich um Werkzeuge handelt), muss man runden.

Es kann sein, dass man keine volle Auslastung will, sondern den Gesamterlös maximieren will. Wenn man für eine ME des Produkts P_k den Preis c_k Euro erzielt, erhält man eine **lineare Optimierungsaufgabe**

$$c_1x_1 + c_2x_2 = \text{Max!}$$

unter der Nebenbedingung

$$a_{11}x_1 + a_{12}x_2 \leq b_1, \quad a_{21}x_1 + a_{22}x_2 \leq b_2$$

und $x_1 \geq 0, x_2 \geq 0$. Meist kommt noch eine Ganzzahligkeitsforderung dazu, wenn die Produkte „unteilbar“ sind, wie z.B. Literpackungen.

Ganz offensichtlich kann man das LGS sofort auf m Maschinen und n Produkte erweitern, so dass man (3.8) erhält. Ist $n > m$, hat man meist noch genügend Spiel, um z.B. den Gesamterlös zu maximieren. Ist $n < m$, wird i.A. eine volle Auslastung nicht möglich sein.

Auch wenn keine Gleichungen gefordert werden, so gibt die Matrix A folgendermaßen einen Sinn: Jedem „Produktionsplan“ \mathbf{x} kann man den „Zeitbedarfsvektor“ \mathbf{b} mit den Komponenten

$$b_j := a_{j1}x_1 + \cdots + a_{jn}x_n, j = 1, 2, \dots, m$$

zuordnen.

Nahrungsmittel und Nährstoffe

Wenn man gewisse Mengeneinheiten von Nahrungsmitteln (Kartoffeln, Milch, Birnen, Möhren, etc.) in einen „Warenkorb“ packt, so enthält dieser damit gewisse Mengeneinheiten von Nährstoffen (Vitamin C, Eiweiß, diverse Fette, etc.). Will man wissen, wieviel von jedem Nährstoff insgesamt im Warenkorb liegt, muss man für jedes Nahrungsmittel wissen, wieviel es von jedem Nährstoff enthält. Genauer:

Es gebe n Nahrungsmittel und m Nährstoffe. Der Warenkorb enthalte x_k Mengeneinheiten (Stückzahl, Liter, kg,...) des k -ten Nahrungsmittels, $k = 1, 2, \dots, n$. Eine ME des k -ten Nahrungsmittels enthalte a_{jk} ME des j -ten Nährstoffs. Dann enthält der Warenkorb insgesamt

$$b_j := a_{j1}x_1 + a_{j2}x_2 + \cdots + a_{jn}x_n$$

ME des j -ten Nährstoffs. Die $m \times n$ -Matrix $A = (a_{jk})_{j,k=1,2,\dots,m,k=1,2,\dots,n}$ kann als rechteckige „Nährstofftabelle“ aufgefasst werden, deren Zeilen die Nährstoffe und deren Spalten die Nahrungsmittel wiedergeben.

A definiert eine *lineare Abbildung* (s. Kap. 3.3) $\mathcal{A} : \mathbb{R}^n \rightarrow \mathbb{R}^m$, die jedem durch $\mathbf{x} \in \mathbb{R}^n$ charakterisierten Warenkorb den zugehörigen „Nährstoffvektor“ \mathbf{b} zuordnet.

3.2.4 Elementare Zeilenumformungen

Vertauscht man zwei Gleichungen in (3.8), so ändert sich nichts an Lösungen \mathbf{x} . Desgleichen, wenn man eine Gleichung mit einem Skalar $\lambda \neq 0$ „durchmultipliziert“ oder wenn man eine mit einem Skalar $\lambda \in \mathbb{R}$ multiplizierte Gleichung zu einer anderen Gleichung addiert. Solche Umformungen nennt man **elementare Zeilenumformungen**, weil — alleine auf A bezogen — das neu entstandene LGS eine Koeffizientenmatrix hat, die aus A durch Zeilentausch, durch Ersetzung der j -ten Zeile \mathbf{a}^j durch $\lambda\mathbf{a}^j$ bzw. durch Ersetzung von \mathbf{a}^j durch $\mathbf{a}^j + \lambda\mathbf{a}^i$ hervorgeht. Bezogen auf das gesamte LGS erstrecken sich die Elementarumformungen auch auf die rechte Seite. Dies kann man sinnvollerweise als weitere Spalte an die Matrix A anfügen (man spricht dann von einer **erweiterten Matrix** $\tilde{A} := (A|\mathbf{b})$) und dann entsprechende Umformungen vornehmen. Da diese nichts an den möglichen Lösungen ändert, kann man sich fragen, in welche Gestalt man das LGS bringen sollte, damit man die Lösung(en) ablesen kann.

3.2.5 Gauß'sches Eliminationsverfahren

Die Idee für den wichtigsten Fall $m = n$ besteht darin, das LGS durch Elementarumformungen in ein **gestaffeltes** LGS zu überführen. Hierunter versteht man ein LGS der Form

$$\begin{aligned}
 a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n &= b_1 \\
 a_{22}x_2 + \dots + a_{2n}x_n &= b_2 \\
 &\vdots \\
 a_{nn}x_n &= b_n
 \end{aligned}$$

Dieses kann man sukzessiv von „unten nach oben“ lösen: aus der letzten Gleichung erhält man x_n (sofern $a_{nn} \neq 0$), setzt diesen Wert in die vorletzte Gleichung ein und löst diese nach x_{n-1} auf (wenn $a_{n-1,n-1} \neq 0$) und fährt so fort, bis man bei der ersten Gleichung angelangt ist. Hier setzt man alle bisher gefundenen Werte x_n, x_{n-1}, \dots, x_2 ein und löst nach x_1 auf (wenn $a_{11} \neq 0$).

Als Beispiel betrachten wir das Produktionsproblem

$$2x_1 + 2x_2 = 32 \quad (3.12)$$

$$x_1 + 3x_2 = 24 \quad (3.13)$$

Nun löst man die erste Gleichung nach x_1 auf und setzt in die zweite Gleichung ein, also $x_1 = 16 - x_2$, eingesetzt in die zweite Gleichung führt auf $16 - x_2 + 3x_2 = 24$, also auf $2x_2 = 8$. Wir erhalten insgesamt das gestaffelte System

$$2x_1 + 2x_2 = 32$$

$$2x_2 = 8$$

das wir von unten nach oben lösen können: $x_2 = 4$ löst die zweite Gleichung. Eingesetzt in die erste ergibt dies $x_1 = 12$. Dieses Vorgehen zeigt deutlich einen **Eliminationsschritt** (x_1 wird eliminiert), nur erkennt man noch nicht direkt eine Elementarumformung. Das geht so: Man addiert die mit -0.5 multiplizierte erste Gleichung in (3.12) zur zweiten.

Ein zweites Beispiel mit $m = n = 3$ ist das folgende (auch hier gibt es ganzzahlige Lösungen):

$$x_1 + x_2 + x_3 = 6$$

$$x_1 + 2x_2 + 3x_3 = 13$$

$$-x_1 + 3x_2 + 2x_3 = 12$$

Die zugehörige (erweiterte) Matrix lautet

$$\tilde{A} = \begin{pmatrix} 1 & 1 & 1 & 6 \\ 1 & 2 & 3 & 13 \\ -1 & 3 & 2 & 12 \end{pmatrix}$$

Nun eliminieren wir x_1 aus der ersten Gleichung, indem wir diese mit -1 multiplizieren und zur zweiten addieren, bzw. sie unverändert zur dritten dividieren. Das neue LGS lautet

$$\begin{aligned}x_1 + x_2 + x_3 &= 6 \\x_2 + 2x_3 &= 7 \\4x_2 + 3x_3 &= 18\end{aligned}$$

Aus der erweiterten Matrix wird

$$\tilde{A} = \begin{pmatrix} 1 & 1 & 1 & 6 \\ 0 & 1 & 2 & 7 \\ 0 & 4 & 3 & 18 \end{pmatrix}$$

Nun multiplizieren wir die zweite (neue) Gleichung mit -4 und addieren sie zur letzten:

$$\begin{aligned}x_1 + x_2 + x_3 &= 6 \\x_2 + 2x_3 &= 7 \\-5x_3 &= -10\end{aligned}$$

Aus der erweiterten Matrix wird

$$\tilde{A} = \begin{pmatrix} 1 & 1 & 1 & 6 \\ 0 & 1 & 2 & 7 \\ 0 & 0 & -5 & -10 \end{pmatrix}$$

Jetzt kann man von unten nach oben auflösen: $x_3 = 2, x_2 = 7 - 2x_3 = 3, x_1 = 6 - x_2 - x_3 = 1$.

Wenn man schon eine Matrizenmultiplikation eingeführt hat, kann man die Elementarumformungen auch als $A \rightarrow L \cdot A$ erkennen, wobei L eine spezielle untere Dreiecksmatrix ist mit lauter Einsen in der Diagonalen.

3.3 Lineare Abbildungen und Matrizen

Betrachtet man das LGS (3.11), so erkennt man einen funktionalen Zusammenhang zwischen einem Vektor $\mathbf{x} = (x_1, x_2)$ und der linken Seite des LGS. Es ist naheliegend, eine Abbildung zu betrachten, die $\mathbf{x} \in \mathbb{R}^2$ den Vektor $\mathbf{y} := (a_{11}x_1 + a_{12}x_2, a_{21}x_1 + a_{22}x_2) \in \mathbb{R}^2$ zuordnet. Diese Abbildung des \mathbb{R}^2 in sich wird durch die 2×2 -Matrix

$$A := \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$$

bestimmt. Um die enge Beziehung zwischen Matrix und Abbildung zu betonen, bezeichnen wir die Abbildung mit $\mathcal{A} : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ und benutzen die Spaltenvektorform:

$$\mathcal{A} : \mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \mapsto \begin{pmatrix} a_{11}x_1 + a_{12}x_2 \\ a_{21}x_1 + a_{22}x_2 \end{pmatrix} \quad (3.14)$$

Die Abbildung \mathcal{A} heißt **lineare Abbildung** und erfüllt die Rechenregeln (für beliebige Vektoren $\mathbf{u}, \mathbf{v} \in \mathbb{R}^2$ und Skalare $\lambda \in \mathbb{R}$):

$$\mathcal{A}(\mathbf{u} + \mathbf{v}) = \mathcal{A}(\mathbf{u}) + \mathcal{A}(\mathbf{v}) \quad (3.15)$$

$$\mathcal{A}(\lambda\mathbf{u}) = \lambda\mathcal{A}(\mathbf{u}) \quad (3.16)$$

3.3.1 Linearität im Nahrungsmittel-Nährstoff-Beispiel

Wir haben in Kap. 3.2.3 einen funktionalen *linearen* Zusammenhang zwischen dem „Warenkorb-Vektor“ $\mathbf{x} \in \mathbb{R}^n$ und dem „Nährstoffvektor“ $\mathbf{y} \in \mathbb{R}^m$ der Form $\mathbf{y} = A\mathbf{x}$ angenommen. Auf einen Nährstoff mit y ME und ein Nahrungsmittel mit x ME bezogen, bedeutet dies die Annahme, dass die Menge y eines in einem Nahrungsmittel enthaltenen Nährstoffes *proportional* zu der ME x des Nahrungsmittels ist, dass es also eine *Konstante* $a > 0$ gibt mit $y = f(x) := a \cdot x$. Man sagt auch, dass y *linear mit x wächst* oder dass die Funktion $f(x) = ax$ ein *lineares Wachstum* beschreibt. Die in der Matrix A auftretenden Elemente sind gerade solche Proportionalitätskonstanten.

Diese mit der Proportionalität verbundene Linearität ist eine Annahme, die keineswegs richtig sein muss. Z.B. enthält eine besonders große Kartoffel nicht immer eine dem Gewicht entsprechende größere Menge an Vitaminen oder Mineralien, weil das höhere Gewicht vielleicht auf einen höheren Wassergehalt zurückzuführen ist.

Auch beim Produktionsmodell kann man die angenommene Linearität (dass die benötigte Zeit einer Maschine proportional zur der zu produzierenden Menge ist) bezweifeln. Aber sehr häufig ist die angenommene Linearität eine gute Annäherung an die Realität.

3.3.2 Lineare Gleichungssysteme und Lineare Abbildungen

Eine Verallgemeinerung auf das LGS (3.8),

$$a_{j1}x_1 + a_{j2}x_2 + \cdots + a_{jn}x_n = b_j, j = 1, 2, \dots, m, \quad (3.17)$$

mit der $m \times n$ -Koeffizientenmatrix

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

ist offensichtlich. Letztere definiert eine Abbildung

$$\mathcal{A} : \mathbb{R}^n \rightarrow \mathbb{R}^m, \quad \mathbf{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \mapsto \begin{pmatrix} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n \end{pmatrix}, \quad (3.18)$$

die ebenfalls die Rechengesetze (3.15) erfüllt.

Das LGS lässt sich jetzt so formulieren: Gegeben sei eine $m \times n$ -Matrix A mit der zugehörigen linearen Abbildung $\mathcal{A} : \mathbb{R}^n \rightarrow \mathbb{R}^m$ sowie ein Vektor $\mathbf{b} \in \mathbb{R}^m$. Gesucht ist ein Urbild \mathbf{x} von \mathbf{b} unter \mathcal{A} , also ein $\mathbf{x} \in \mathbb{R}^n$ mit $\mathcal{A}(\mathbf{x}) = \mathbf{b}$.

Schreibweisen

Im Folgenden wird nicht zwischen der *Matrix* A und der linearen Abbildung \mathcal{A} unterschieden. Statt $\mathcal{A}(\mathbf{x})$ schreibt man kurz nur $A\mathbf{x}$ und merkt sich die Rechenregeln in Form von $A(\mathbf{u} + \mathbf{v}) = A\mathbf{u} + A\mathbf{v}$, $A(\lambda\mathbf{u}) = \lambda A\mathbf{u}$, wobei immer dann eine runde Klammer gesetzt wird, wenn sonst Missverständnisse möglich sind.

Ein zugehöriges LGS wird als $A\mathbf{x} = \mathbf{b}$ notiert - wie schon oben gesagt.

3.3.3 Linearität

Die eben genannten Rechenregeln $A(\mathbf{u} + \mathbf{v}) = A\mathbf{u} + A\mathbf{v}$, $A(\lambda\mathbf{u}) = \lambda A\mathbf{u}$ stehen für die *Linearität* der Abbildung. Warum wird diese Eigenschaft so betont?³

Als erstes ist das *Superpositionsprinzip* zu nennen: Wenn man zwei LGSe mit derselben Matrix und zwei verschiedenen rechten Seiten $\mathbf{b}_1, \mathbf{b}_2$ lösen kann, so kann sofort auch das LGS $A\mathbf{x} = \mathbf{b}_1 + \mathbf{b}_2$ gelöst werden, und zwar durch die Summe der beiden Lösungen $\mathbf{x}_1, \mathbf{x}_2$ zu $A\mathbf{x}_1 = \mathbf{b}_1, A\mathbf{x}_2 = \mathbf{b}_2$. Auch eine Abänderung der rechten Seite \mathbf{b} von $A\mathbf{x} = \mathbf{b}$ zu $\lambda\mathbf{b}$ mit einem Skalar $\lambda \in \mathbb{R}$ führt auf die Lösung $\lambda\mathbf{x}$ an Stelle von \mathbf{x} . In gewisser Weise ist davon schon bei dem Eliminationsverfahren zur Lösung von LGSen Gebrauch gemacht worden.

Lösungstheorie

An dieses Konzept der Linearität knüpft die *Lösungstheorie* für allgemeine LGSe an, die ich nur skizzieren möchte. Dabei nennt man das LGS **homogen**, wenn $\mathbf{b} = \mathbf{o}$ und sonst **inhomogen**. Das homogene LGS hat stets die Lösung $\mathbf{x} = \mathbf{o}$, die man als **triviale Lösung** des homogenen Systems bezeichnet. Alle Lösungen des homogenen LGS bilden einen Vektorraum, einen sog. *Unterraum* des \mathbb{R}^n , den man **Kern** von A nennt. Dieser hat eine gewisse *Dimension* — das ist die maximale Anzahl linear unabhängiger Lösungen. Dem steht das **Bild** von A gegenüber, also

$$\text{Bild}(A) := \{\mathbf{y} \in \mathbb{R}^m : \exists \mathbf{x} \in \mathbb{R}^n : A\mathbf{x} = \mathbf{y}\}.$$

³Dieses Konzept durchzieht die gesamte Mathematik!

Auch $Bild(A)$ ist ein Vektorraum — ein Unterraum von \mathbb{R}^m . Dessen Dimension nennt man den **Rang** der Matrix A . Es gilt die Dimensionsformel

$$n = \dim(Kern(A)) + \dim(Bild(A)).$$

Nur für einen „trivialen“ Kern kann das LGS eine eindeutige Lösung besitzen: Aus $A\mathbf{x} = \mathbf{b}$ und $A\mathbf{v} = \mathbf{0}$ folgt, dass auch $A(\mathbf{x} + \mathbf{v}) = \mathbf{b}$. Nur, wenn $Bild(A) = \mathbb{R}^m$, was schon durch $Rang(A) = m$ garantiert ist, besitzt das LGS für jede rechte Seite eine Lösung. Insgesamt gibt es nur dann für alle rechten Seiten \mathbf{b} eine eindeutige Lösung \mathbf{x} von $A\mathbf{x} = \mathbf{b}$, wenn $m = n$ (Anzahl der Gleichungen gleich Anzahl der Unbekannten) und wenn der Kern von A nur aus dem Nullvektor besteht. In diesem Fall kann man eine *Determinante* $\Delta(A)$ definieren, die diese Eigenschaft durch $\Delta(A) \neq 0$ garantiert.

Diese letzten Aussagen hängen mit der *Injektivität*, *Surjektivität* und *Bijektivität* von linearen Abbildungen zusammen: Da $\mathcal{A}(\mathbf{x})$ eine Linearkombination der Spalten von A ist, ist die Injektivität gegeben, wenn die n Spalten $\mathbf{a}_i \in \mathbb{R}^m, i = 1, 2, \dots, n$ linear unabhängig sind, was nur für $m \geq n$ sein kann. \mathcal{A} ist surjektiv, wenn sich jedes $\mathbf{b} \in \mathbb{R}^m$ linear durch die Spalten $\mathbf{a}_i \in \mathbb{R}^m, i = 1, 2, \dots, n$ kombinieren lässt. Etwas schwieriger einzusehen ist, dass dies nur für $m \leq n$ sein kann. Bijektivität kann nur für $m = n$ vorliegen und wird durch $\Delta(A) \neq 0$ garantiert.

3.3.4 Geometrische Transformationen und lineare Abbildungen

In der Computergrafik will man gewisse Objekte „transformieren“. Dabei ist ein Objekt eine Menge M von Punkten des \mathbb{R}^3 (3D) oder des \mathbb{R}^2 (2D), für deren Bild $\mathcal{A}(M)$ (als Bildmenge) man sich interessiert.

Die einfachsten geometrischen Transformationen sind Drehungen und Spiegelungen von ebenen Objekten im \mathbb{R}^2 .

Drehungen: Zu einem Drehwinkel α mit $0 \leq \alpha < 2\pi$ gehört eine Abbildung $\mathcal{R}_\alpha : \mathbb{R}^2 \rightarrow \mathbb{R}^2$, die einen Punkt $P = (x, y)$ um eine Drehachse, die durch den Ursprung senkrecht zur (x, y) -Ebene verläuft, um den Winkel α gegen den Uhrzeigersinn dreht. Man kann sich leicht überzeugen, dass \mathcal{R}_α eine lineare Abbildung ist. Da für allgemeine lineare Abbildungen $\mathcal{A} : \mathbb{R}^n \rightarrow \mathbb{R}^m$ ja $\mathcal{A}(\mathbf{x}) = \sum_{i=1}^n x_i \mathbf{a}_i$ mit den Spalten $\mathbf{a}_i \in \mathbb{R}^m, i = 1, 2, \dots, n$ der $(m \times n)$ -Matrix A gilt, kann man die beiden Spalten der zu \mathcal{R}_α gehörenden Drehmatrix R_α leicht berechnen, und es ergibt sich

$$R_\alpha = \begin{pmatrix} \cos(\alpha) & -\sin(\alpha) \\ \sin(\alpha) & \cos(\alpha) \end{pmatrix}.$$

Beachten Sie, dass für den Drehwinkel $\alpha := 0$ die Identität \mathcal{R}_0 ergibt, welche wiederum durch die *Einheitsmatrix* dargestellt wird.

Es ist jetzt außerordentlich nützlich und ästhetisch, dass aus der offensichtlich richtigen Aussage $\mathcal{R}_\alpha \circ \mathcal{R}_\beta = \mathcal{R}_{\alpha+\beta}$ die Additionstheoreme für Sinus und Kosinus folgen, weil

$$\begin{pmatrix} \cos(\alpha) & -\sin(\alpha) \\ \sin(\alpha) & \cos(\alpha) \end{pmatrix} \cdot \begin{pmatrix} \cos(\beta) & -\sin(\beta) \\ \sin(\beta) & \cos(\beta) \end{pmatrix} = \begin{pmatrix} \cos(\alpha + \beta) & -\sin(\alpha + \beta) \\ \sin(\alpha + \beta) & \cos(\alpha + \beta) \end{pmatrix}$$

sein muss, sich aber durch Matrizenmultiplikation (siehe Kap. 3.3.7, Beispiel 1)

$$\begin{pmatrix} \cos(\alpha) & -\sin(\alpha) \\ \sin(\alpha) & \cos(\alpha) \end{pmatrix} \cdot \begin{pmatrix} \cos(\beta) & -\sin(\beta) \\ \sin(\beta) & \cos(\beta) \end{pmatrix} = \\ \begin{pmatrix} \cos(\alpha)\cos(\beta) - \sin(\alpha)\sin(\beta) & -(\cos(\alpha)\sin(\beta) + \sin(\alpha)\cos(\beta)) \\ \cos(\alpha)\sin(\beta) + \sin(\alpha)\cos(\beta) & \cos(\alpha)\cos(\beta) - \sin(\alpha)\sin(\beta) \end{pmatrix}$$

berechnet.

Spiegelungen:

Wenn man an einer Geraden durch den Ursprung mit Neigungswinkel ϕ spiegelt, so wurde in Übungsaufgabe 9 (in der jetzigen Sprechweise) gezeigt, dass die zugehörige Abbildung linear ist und dass sich diese durch eine „Spiegelungsmatrix“

$$S_\phi = \begin{pmatrix} \cos(2\phi) & \sin(2\phi) \\ \sin(2\phi) & -\cos(2\phi) \end{pmatrix}.$$

darstellen lässt.

Interessant ist das Zusammenspiel von Drehungen und Spiegelungen bei Verkettungen. Verkettet man zwei Spiegelungen, erhält man eine Drehung. Verkettet man eine Spiegelung mit einer Drehung, erhält man eine Spiegelung. Hier kann man auch Symmetriegruppen wiederfinden, siehe Vorlesung. So sind die vier Gruppen $(\mathbf{Z}_4, +_4)$, $(\{R_\alpha : \alpha = 0, \pi/2, \pi, 3\pi/2\}, \cdot)$ mit der Matrizenmultiplikation \cdot , $(\{\mathcal{R}_\alpha : \alpha = 0, \pi/2, \pi, 3\pi/2\}, \circ)$ mit der Verkettung \circ und $(\{1, i, -1, -i\}, \cdot)$ mit der komplexen Multiplikation i.W. die gleichen Gruppen, sie sind alle „isomorph“.

Drehungen (um irgendwelche Drehachsen und Drehwinkel) gibt es auch im \mathbb{R}^3 , siehe **Drehmatrizen** (Wikipedia), wo Sie die (3×3) -Drehmatrizen finden können.

Bei Spiegelungen im \mathbb{R}^3 spiegelt man an Ebenen durch den Ursprung.

Ganz allgemein ist schon im \mathbb{R}^2 die Frage berechtigt, wie eine durch eine Matrix

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$$

gegebene lineare Abbildung geometrisch zu interpretieren ist.

Dies ist relativ einfach bei *Diagonalmatrizen*

$$D := \begin{pmatrix} d_1 & 0 \\ 0 & d_2 \end{pmatrix}$$

mit *Diagonalelementen* $d_j > 0, j = 1, 2$. Bei der zugehörigen Abbildung handelt es sich um eine Streckung in x - und y -Richtung.



Abbildung 3.1: Original

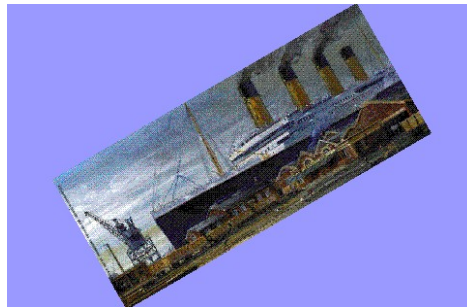


Abbildung 3.2: Rotation

Die durch die Matrix

$$S_x := \begin{pmatrix} 1 & a \\ 0 & 1 \end{pmatrix}$$

mit $a > 0$ gegebene Abbildung kann man als *Scherung* in x -Richtung interpretieren, während die zur Matrix

$$S_y := \begin{pmatrix} 1 & 0 \\ b & 1 \end{pmatrix}$$

gehörende lineare Abbildung als *Scherung* in y -Richtung gedeutet werden kann.

Am besten macht man sich dies klar, wenn man die Punktmenge

$$Q := \{(x, y) \in \mathbb{R}^2 : 0 \leq x \leq 1, 0 \leq y \leq 1\} \quad (3.19)$$

— das Einheitsquadrat — abbildet.

In den Abbildungen 3.1-3.4 sehen Sie die geometrische Interpretation der verschiedenen Abbildungen:

Ein wenig allgemeiner als lineare Abbildungen sind **affine Abbildungen**. Dies sind Abbildungen der Form $\mathbf{x} \mapsto \mathcal{A}(\mathbf{x}) + \mathbf{a}$ mit einer linearen Abbildung \mathcal{A} . Es handelt sich hierbei also um eine Verkettung einer linearen Abbildung und einer Translation $T_{\mathbf{a}} \circ \mathcal{A}$. Auf der Webseite [Affine Abbildungen](#) (Uni Flensburg) werden diese Abbildungen visualisiert. Natürlich sind lineare Abbildungen Spezialfälle affiner Abbildungen ($\mathbf{a} = \mathbf{o}$).

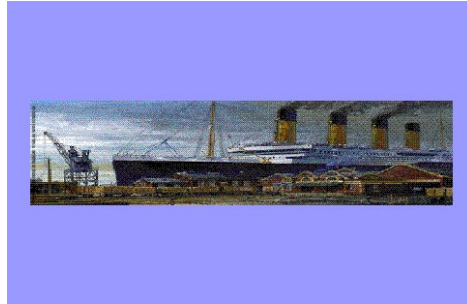


Abbildung 3.3: Dehnung

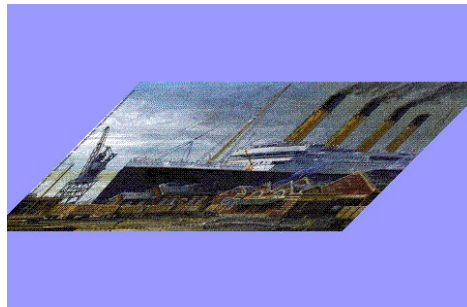


Abbildung 3.4: Scherung

3.3.5 Ausnutzung der Linearität

Jetzt betrachten wir wieder eine allgemeine lineare Abbildung $\mathcal{A} : \mathbb{R}^n \rightarrow \mathbb{R}^m$. Wir wissen, dass eine Matrix

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

eine solche lineare Abbildung definiert:

Bezeichnet man wieder die Spaltenvektoren von A mit $\mathbf{a}_j, j = 1, 2, \dots, n$, so ist durch $\mathcal{A}(\mathbf{x}) := A\mathbf{x} = x_1\mathbf{a}_1 + x_2\mathbf{a}_2 + \cdots + x_n\mathbf{a}_n = \sum_{j=1}^n x_j\mathbf{a}_j$ eine lineare Abbildung \mathcal{A} schon festgelegt, wenn $\mathbf{a}_j := \mathcal{A}(\mathbf{e}_j), j = 1, 2, \dots, n$, für die n Einheitsvektoren \mathbf{e}_j des \mathbb{R}^n gegeben sind, $j = 1, 2, \dots, n$.

Ist eine lineare Abbildung \mathcal{A} gegeben, so gehört dazu stets eine Matrix, wenn man die Basis der Einheitsvektoren zu Grunde legt: Man setze nur $\mathbf{a}_j := \mathcal{A}(\mathbf{e}_j), j = 1, 2, \dots, n$, für die n Einheitsvektoren \mathbf{e}_j des \mathbb{R}^n und bilde die zugehörige Matrix A mit den Spalten \mathbf{a}_j .

Ein wichtiger, bisher kaum behandelte Spezialfall ist $m = n$ mit der Identität \mathcal{A} . Die zugehörige

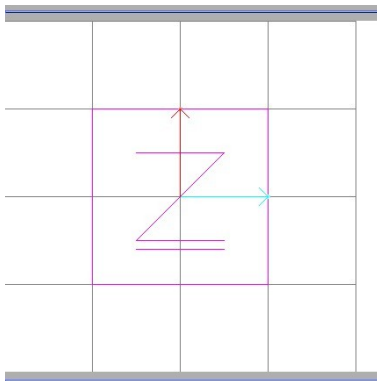


Abbildung 3.5: Ausgangsbild: Buchstabe Z

Matrix A ist die $n \times n$ -**Einheitsmatrix**

$$E = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix}$$

Als Nebenprodukt können wir für den ebenen Fall $m = 2 = n$ die Frage nach $A(Q)$ mit dem Einheitsquadrat Q in (3.19) leicht beantworten: $A(Q)$ ist das Parallelogramm zu den Kantenvektoren $\mathbf{a}_1, \mathbf{a}_2$.

3.3.6 Internet

Auf der Webseite [Matrix Simulation](#) (Paul Falstad) kann man den Buchstaben Z einer geometrischen Transformation zu einer 2×2 -Matrix unterziehen, z.B. einer Rotation um den Winkel $\pi/6$, einer Spiegelung an einer der Achsen, etc.

Siehe Abb. 3.5, 3.6, 3.7

3.3.7 Verkettung von linearen Abbildungen – Multiplikation von Matrizen

Ist $\mathcal{A} : \mathbb{R}^n \rightarrow \mathbb{R}^m$ eine lineare Abbildung, so kann man sie mittels $\mathcal{C} := \mathcal{A} \circ \mathcal{B}$ mit einer weiteren linearen Abbildung $\mathcal{B} : \mathbb{R}^p \rightarrow \mathbb{R}^n$, wenn die Bilder von \mathcal{B} also Abmessungen haben, die für \mathcal{A} geeignet sind. Das Ergebnis ist eine lineare Abbildung $\mathcal{C} : \mathbb{R}^p \rightarrow \mathbb{R}^m$. Daher kann man die Frage stellen, wie sich die zugehörigen Matrizen A, B und C zueinander verhalten. Das Ergebnis wird durch

$$C = A \cdot B$$

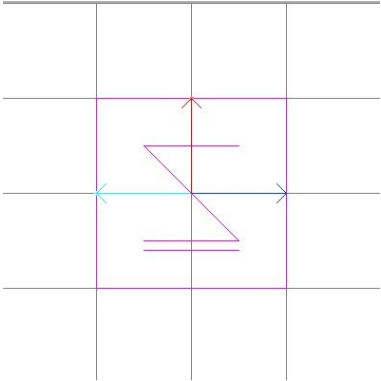


Abbildung 3.6: An der y -Achse gespiegeltes Z

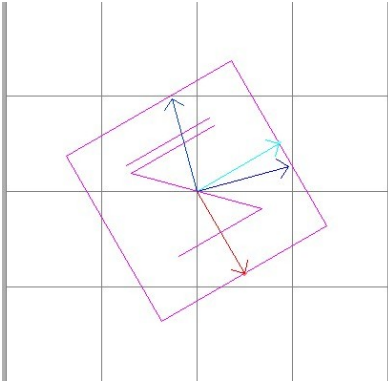


Abbildung 3.7: An der x -Achse gespiegeltes und danach gedrehtes Z

ausgedrückt, wobei „ \cdot “ eine *Matrizenmultiplikation* beschreibt. A ist eine $m \times n$ -Matrix, B ist eine $n \times p$ -Matrix und C muss eine $m \times p$ -Matrix werden. Die Herleitung von „Formeln“ dieses Matrizenprodukts ist nicht schwer, man muss nur aufpassen, von wo bis wo die Indizes laufen: Die p Spalten $\mathbf{c}_k \in \mathbb{R}^m, k = 1, 2, \dots, p$ von C müssen ja gerade $\mathbf{c}_k = \mathcal{C}(\mathbf{e}_k) = \mathcal{A}(\mathcal{B}(\mathbf{e}_k))$ genügen. $\mathbf{b}_k = \mathcal{B}(\mathbf{e}_k) \in \mathbb{R}^n, k = 1, 2, \dots, p$ sind die p Spalten von B , d.h.

$$\mathbf{b}_k = \begin{pmatrix} b_{1k} \\ \vdots \\ b_{nk} \end{pmatrix}.$$

Also gilt $\mathbf{c}_k = \mathcal{A}(\mathbf{b}_k), k = 1, 2, \dots, p$, bzw.

$$\mathbf{c}_k = \begin{pmatrix} c_{1k} \\ \vdots \\ c_{mk} \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} b_{1k} \\ \vdots \\ b_{nk} \end{pmatrix}.$$

Dies führt auf

$$c_{jk} = \mathbf{a}^j \cdot \mathbf{b}_k = a_{j1}b_{1k} + a_{j2}b_{2k} + \cdots + a_{jn}b_{nk}, j = 1, 2, \dots, m, k = 1, 2, \dots, p$$

mit dem j -ten Zeilenvektor \mathbf{a}^j von A .

Beispiel 1 für $m = n = p = 2$ (siehe auch Kap. 3.3.4). Es geht also um das Produkt von 2×2 -Matrizen. Die beiden Abbildungen \mathcal{A} und \mathcal{B} seien beides Drehungen, einmal um den Winkel φ , also

$$A := \begin{pmatrix} \cos(\varphi) & -\sin(\varphi) \\ \sin(\varphi) & \cos(\varphi) \end{pmatrix}$$

und um den Winkel ψ , also

$$A := \begin{pmatrix} \cos(\psi) & -\sin(\psi) \\ \sin(\psi) & \cos(\psi) \end{pmatrix}$$

Nach den Regeln der Matrizenmultiplikation gilt für

$$C = \begin{pmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{pmatrix}$$

$$c_{11} = a_{11}b_{11} + a_{12}b_{21}, c_{12} = a_{11}b_{12} + a_{12}b_{22},$$

$$c_{21} = a_{21}b_{11} + a_{22}b_{21}, c_{22} = a_{21}b_{12} + a_{22}b_{22},$$

also wegen der Additionstheoreme

$$c_{11} = \cos(\varphi)\cos(\psi) - \sin(\varphi)\sin(\psi) = \cos(\varphi + \psi),$$

$$c_{12} = -\cos(\varphi)\sin(\psi) - \sin(\varphi)\cos(\psi) = -\sin(\varphi + \psi),$$

$$c_{21} = \sin(\varphi)\cos(\psi) + \cos(\varphi)\sin(\psi) = \sin(\varphi + \psi),$$

$$c_{22} = -\sin(\varphi)\sin(\psi) + \cos(\varphi)\cos(\psi) = \cos(\varphi + \psi).$$

Wie erwartet, ist C wieder eine „Drehmatrix“ zum Winkel $\varphi + \psi$.

Beispiel 2:

Sei $\mathcal{A} : \mathbb{R}^3 \rightarrow \mathbb{R}^4, (x, y, z) \mapsto (x + y, y - 3z, 2x - 6z, x - y)$ und $\mathcal{B} : \mathbb{R}^2 \rightarrow \mathbb{R}^3, (u, v) \mapsto (2u - v, 3v, 4u)$. Die zugehörigen Matrizen sind:

$$A = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & -3 \\ 2 & 0 & -6 \\ 1 & -1 & 0 \end{pmatrix} \quad B := \begin{pmatrix} 2 & -1 \\ 0 & 3 \\ 4 & 0 \end{pmatrix}$$

$\mathcal{C} := \mathcal{A} \circ \mathcal{B}$ bildet den \mathbb{R}^2 in den \mathbb{R}^4 ab und hat die Abbildungsvorschrift (Setze $x := 2u - v, y := 3v, z := 4u$)

$$\mathcal{C} : (u, v) \mapsto ((2u - v) + (3v), 3v - 3(4u), 2(2u - v) - 6(4u), (2u - v) - (3v)) = (2u + 2v, -12u + 3v, -20u - 2v, 2u - 3v)$$

die zugehörige Matrix ist

$$C := \begin{pmatrix} 2 & 2 \\ -12 & 3 \\ -20 & -2 \\ 2 & -4 \end{pmatrix}$$

die sich offensichtlich auch als Matrizenprodukt $C = A \cdot B$ schreiben lässt.

Beispiel 3: Das Skalarprodukt $\mathbf{a} \cdot \mathbf{b} = a_1b_1 + a_2b_2 + \dots + a_nb_n$ kann man auch als Produkt einer $(1 \times n)$ -Matrix mit einer $(n \times 1)$ Matrix mit dem ergebnis einer (1×1) -Matrix auffassen:

$$(a_1 \quad a_2 \quad \dots \quad a_n) \cdot \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix} = a_1b_1 + a_2b_2 + \dots + a_nb_n.$$

Beispiel 4:

Sei A eine $(m \times n)$ -Matrix und $\mathbf{x} \in \mathbb{R}^n$ ein Vektor aufgefasst als $(n \times 1)$ -Matrix, also als Spaltenvektor. Dann ist das Matrizenprodukt $\mathbf{b} := A \cdot \mathbf{x}$ eine $(m \times 1)$ -Matrix ist und gleich $\mathbf{b} = A\mathbf{x}$, wobei $A\mathbf{x} := \mathcal{A}(\mathbf{x})$ mit der zu A gehörenden linearen Abbildung ist.

3.4 Zentralprojektion

Wir schließen dieses Kapitel mit einer Betrachtung einer in der Computergrafik wichtigen Abbildung, die auch in Übungs-Aufgabe 39 behandelt wurde.

Bei der **Zentralprojektion** wird mit Hilfe eines *Auges* $A \in \mathbb{R}^3$ ein Punkt $P \in \mathbb{R}^3$ auf eine *Projektionsebene* \mathcal{E} projiziert, indem man den Schnittpunkt S der Ebene \mathcal{E} mit der Verbindungsgeraden \mathcal{G}_{AP} als *Bildpunkt* wählt: Die Zentralprojektion ist also eine Abbildung namens Z , die einem Punkt $P := (x_1, x_2, x_3)$ eben diesen Schnittpunkt S als Bildpunkt zuordnet. Offensichtlich gibt es einen funktionalen Zusammenhang zwischen der Abbildung Z und dem Auge A sowie der Projektionsebene \mathcal{E} .

Wir sind frei, ein geeignetes Koordinatensystem zu wählen. Dabei wählen wir die x_1 - x_3 -Ebene ($x_2 = 0$) als Projektionsebene \mathcal{E} und den Punkt $A = (0, d, 0)$, wobei $d > 0$ offensichtlich der Abstand des Auges zur Projektionsebene ist, als „Auge“. Dann hängt die Zentralprojektion Z nur noch von d ab.

Wir suchen jetzt eine Formel für S . Es ist $\mathcal{G}_{AP} = \{A + t(P - A) : t \in \mathbb{R}\}$ eine „natürliche“ Parameterform der Geraden durch Auge und abzubildenden Punkt. Es gilt $A + t(P - A) = (tx_1, d + t(x_2 - d), tx_3)$. Gesucht ist das t , für das $S = A + t(P - A) \in \mathcal{E}$. Dies ist dadurch gegeben, dass die mittlere Komponente von S verschwindet, also durch $d + t(x_2 - d) = 0$ bzw. durch $t = \frac{d}{d-x_2}$ gegeben. Der Bildpunkt ist also

$$S = \left(\frac{dx_1}{d-x_2}, 0, \frac{dx_3}{d-x_2} \right),$$

der nur für $d \neq x_2$ definiert ist. Das ist genau die Bedingung, dass $A \neq P$ und \mathcal{G}_{AP} nicht parallel zu \mathcal{E} verläuft, d.h., dass $P \notin \mathbb{R}_d^3 := \{(x, y, z) \in \mathbb{R}^3 : y = d\}$.

Beispiel: Betrachte die 8 Eckpunkte $P_j, j = 1, 2, \dots, 8$ eines Quaders mit $P_1 := (0, 0, 3), P_2 := (-3, -6, 0), P_3 := (12, -9, 9), P_4 := (-7, -2, 14), P_5 = (9, -15, 6), P_6 = (-10, -8, 11), P_7 = 5, -11, 20)$ und $P_8 = (2, -17, 17)$ und bilde sie mittels Zentralprojektion mit $d = 12$ ab.

Die Bildpunkte (x_1, x_3) lauten $P'_1 = (0, 3), P'_2 = (-2, 0), P'_3 = (6 \frac{6}{7}, 5 \frac{1}{7}), P'_4 = (-6, 12), P'_5 = (4, 2 \frac{2}{3}), P'_6 = (-6, 6 \frac{3}{5}), P'_7 = (2 \frac{14}{23}, 10 \frac{10}{23}), P'_8 = (24/29, 7 \frac{1}{29})$.

Siehe Abb. 3.8.

Die Abbildung Z der Zentralprojektion bildet \mathbb{R}_d^3 nach \mathbb{R}^2 ab, wenn man

$$\mathbf{x} = (x_1, x_2, x_3) \mapsto \left(\frac{dx_1}{d-x_2}, \frac{dx_3}{d-x_2} \right)$$

betrachtet. Diese Abbildung ist zwar nicht linear, aber wenigstens *gebrochen linear*, die rechte Seite ist eine *rationale Funktion*.

Man kann jetzt allerlei Interessantes „entdecken“. Z.B., dass alle vor dem Auge liegenden Halbgeraden $\{P + \lambda \mathbf{v} : \lambda \geq 0\}$ mit einem gemeinsamen Richtungsvektor \mathbf{v} in nicht-parallele Strecken abgebildet werden, die sich in einem nur von \mathbf{v} abhängigen Schnittpunkt schneiden, der den

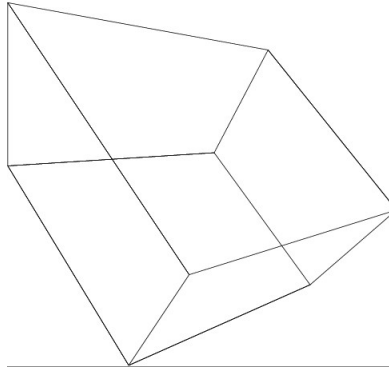


Abbildung 3.8: Zentralprojizierter Quader — Aufgabe 39

„unendlich fernen“ Punkten der Halbgeraden entspricht. Dieser Schnittpunkt heißt **Fluchtpunkt**. Durchläuft man mit \mathbf{v} alle Richtungen (bis auf die, die parallel zur Projektionsebene verlaufen) einer Ebene, so stellt sich heraus, dass alle Fluchtpunkte auf einer Geraden liegen, der sog. **Fluchtgeraden**. Dies zu zeigen, ist gar nicht so schwer, wenn man etwas über Grenzwerte aus der Analysis heranzieht⁴:

Die Zentralprojektion ist also die Abbildung

$$\mathbf{x} = (x, y, z) \mapsto \left(\frac{dx}{d-y}, \frac{dz}{d-y} \right),$$

die jedem Punkt $\mathbf{x} \in \mathbb{R}^3$, der nicht in einer zur Projektionsebene parallelen, durch das „Auge“ $(0, d, 0)$ gehenden Ebene liegt, einen Bildpunkt in der (x, z) -Ebene, also dem \mathbb{R}^2 zuordnet. Nun betrachten wir eine Halbgerade $\{P + \lambda \mathbf{v} : \lambda \geq 0\}$ mit einem Richtungsvektor $\mathbf{v} := (v_1, v_2, v_3)$, einem Punkt $P := (p_1, p_2, p_3)$ und eine gegen *Unendlich divergierende Folge* (λ_n) mit den zugehörigen *Punktfolgen*

$$X_n = (x_n, y_n, z_n) = (p_1 + \lambda_n v_1, p_2 + \lambda_n v_2, p_3 + \lambda_n v_3).$$

Bilden wir X_n durch die Zentralprojektion Z ab, so erhalten wir eine Punktfolge

$$S_n := \left(\frac{dx_n}{d-y_n}, \frac{dz_n}{d-y_n} \right)$$

in der Projektionsebene.

Nun betrachte man die erste Komponente von S_n . Sie lautet wegen $y_n = p_2 + \lambda_n v_2$

$$\frac{dx_n}{d-y_n} = \frac{d(p_1 + \lambda_n v_1)}{d - p_2 - \lambda_n v_2}$$

⁴Das Folgende ist ein Auszug aus meinem Analysis-Skript

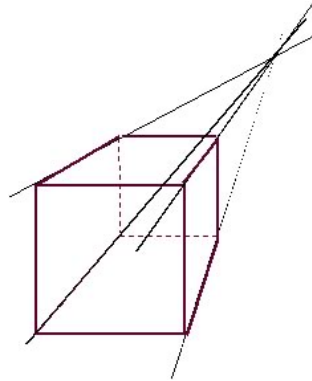


Abbildung 3.9: Fluchtpunkt

und nach Erweiterung mit $\frac{1}{\lambda_n}$

$$d \frac{\frac{p_1}{\lambda_n} + v_1}{\frac{d-p_2}{\lambda_n} - v_2}.$$

Da $\frac{p_1}{\lambda_n}$ und $\frac{d-p_2}{\lambda_n}$ Nullfolgen sind (weil λ_n gegen Unendlich divergiert), konvergiert die erste Komponente von S_n gegen $-d\frac{v_1}{v_2}$. Natürlich nur, wenn $v_2 \neq 0$, d.h., wenn die Gerade \mathcal{G} nicht parallel zur Projektionsebene liegt — eine natürliche Voraussetzung. Entsprechend konvergiert die zweite Komponente von S_n gegen $-d\frac{v_3}{v_2}$. In einem offensichtlichen Sinne konvergiert S_n gegen den **Fluchtpunkt**

$$S := -d\left(\frac{v_1}{v_2}, \frac{v_3}{v_2}\right).$$

Dieser hängt nicht von $P \in \mathcal{G}$ ab und ist daher für alle zu \mathcal{G} parallelen Geraden gleich. Es ist der Punkt, in dem sich alle zueinander parallelen Geraden nach der Projektion „schneiden“. Zu jeder Parallelschar von Geraden gehört also genau ein Fluchtpunkt, siehe Abb. 3.9!

Entsprechend gehört zu jeder Schar paralleler Ebenen eine **Fluchtgerade**. Sie besteht aus allen Fluchtpunkten von in dieser Ebene liegenden Scharen paralleler Geraden.

Man kann den zu einer Geraden gehörenden Fluchtpunkt konstruieren, indem man die durch das Auge A hierzu parallele Gerade mit der Projektionsebene schneidet. Entsprechend ist eine zu einer Ebene gehörende Fluchtgerade durch die Schnittgerade der durch A verlaufenden parallelen Ebene mit der Projektionsebene gegeben.

3.5 Parallelprojektion

Ein Punkt $P := (x, y, z) \in \mathbb{R}^3$ werde auf folgende Weise auf die (x, z) -Ebene „in Richtung $\mathbf{v} := (v_1, v_2, v_3) \neq \mathbf{o}$ projiziert“: Der Bildpunkt $P' := (x', z')$ ist der Schnittpunkt der Geraden durch P mit Richtungsvektor \mathbf{v} . Hierdurch ist für den Fall, dass \mathbf{v} nicht parallel zur (x, z) -Ebene liegt, d.h., dass $v_2 \neq 0$, eine Projektion $\mathcal{P} : \mathbb{R}^3 \rightarrow \mathbb{R}^2, (x, y, z) \mapsto (x', z')$ definiert. Man kann

sich diese Projektion so vorstellen, dass es eine unendlich weit entfernte Lichtquelle gibt, die parallele Strahlen in Richtung \mathbf{v} schickt, etwa wie die Sonne.

Die Gerade durch P mit Richtungsvektor \mathbf{v} hat die Parameterkurzform $\mathbf{x} = (x_1, x_2, x_3) = P + \lambda\mathbf{v}$. Das λ muss so gewählt sein, dass $x_2 = y + \lambda v_2 = 0$, so dass $\lambda = -\frac{y}{v_2}$ folgt. So ergibt sich

$$x' = x - \frac{y}{v_2}v_1, \quad z' = z - \frac{y}{v_2}v_3,$$

also

$$\mathcal{P} : (x, y, z) \mapsto \left(x - \frac{v_1}{v_2}y, z - \frac{v_3}{v_2}y\right)$$

Die zugehörige Matrix lautet

$$P := \begin{pmatrix} 1 & -\frac{v_1}{v_2} & 0 \\ 0 & -\frac{v_3}{v_2} & 1 \end{pmatrix}.$$