Literaturhinweise

Das Angebot von Lehrbüchern zum Thema „Höhere Mathematik“ ist riesig; die hier zusammengestellte kleine Auswahl besteht aus Büchern, die erstens in der Mathematischen Bibliothek zu finden sind (größtenteils im allgemeinen Bestand, teilweise auch in der Lehrbuchsammlung) und die zweitens von der Stoffauswahl her ungefähr dem Inhalt dieser Vorlesung entsprechen – auch wenn sie meist erheblich mehr Stoff enthalten, als in der Vorlesung behandelt werden kann. Am nützlichsten dürften wohl die beiden erstgenannten Bücher **[D]** und **[MV]** sein. Natürlich werden wir auch gelegentlich Stoff aus der Analysis I benötigen; Referenz hierfür ist das Skriptum

[Ba] H.-J. BARTELS: *Analysis I,* Universität Mannheim, 1998/99, `http://www.math.uni-mannheim.de/~bartels/ana1.ps`

Für den Stoff der höheren Mathematik seien genannt

[D] H.J. DIRSCHMID: *Mathematische Grundlagen der Elektrotechnik,* Vieweg, 1992

Etwa fünf Pfund Mathematik, die nicht nur diese Vorlesung mehr als abdecken, sondern für viele Studenten für deren gesamtes Berufsleben ausreichen dürften. Der Schwerpunkt liegt allerdings eindeutig auf dem Gebiet der Analysis; numerische Mathematik und die Statistik sind so gut wie nicht vorhanden.

[MV] K. MEYBERG, P. VACHENAUER: *Höhere Mathematik I,* Springer, 1995

Erster Band eines zweibändigen Werks; enthält die für die *Höhere Mathematik* relevanten Teile der *Analysis I* sowie den gesamten Stoff der Vorlesung *Höhere Mathematik I* (für die *Höhere Mathematik II* werden zusätzlich noch einige Teile des zweiten Bands benötigt), dazu etwas Numerik. Die Darstellung ist recht kompakt mit fast vollständigen Beweisen. Für einige Grundalgorithmen sind Programme angegeben.

[BDH] BRAUCH/DREYER/HAACKE: *Mathematik für Ingenieure Maschinenbau/Elektrotechnik,* Teubner 1977

Dieses Buch richtet sich an Studenten von Fachhochschulen, deckt aber auch einen großen Teil dieser Vorlesung ab; für uns zu knapp ist vor allem die Darstellung der Vektoranalysis. Dafür gibt es Kapitel, die auch Themen wie Statistik, Fehlerrechnung, *usw.* darstellen.

[F] P. FURLAN: *Das gelbe Rechenbuch 1–3,* Verlag Martina Furlan, Dortmund, *o.J.*

Wer vor allem auf Drill und durchgerechnete Beispiele Wert legt, findet hier laut Untertitel „Rechenverfahren der Höheren Mathematik in Einzelschritten erklärt. Mit vielen ausführlich gerechneten Beispielen“. Die drei Bände decken zwar nicht alles ab, aber doch ziemlich viel.

Dazu seien noch einige „Klassiker“ genannt, die seit mehreren Jahrzehnten in immer neuen Auflagen erscheinen und die auch heute noch interessant sind. Es handelt sich um Werke aus meist recht vielen Bänden, wobei selbst der Stoff der Vorlesung *Höhere Mathematik I* je nach Organisation des Gesamtwerks auf bis zu drei Bände verteilt sein kann. Wegen der Vielzahl von Auflagen und Bänden verzichte ich auf die Angabe von Erscheinungsjahren.

[S] W.I. SMIRNOW: *Lehrbuch der Höheren Mathematik,* VEB Deutscher Verlag der Wissenschaften

Ein Klassiker, nach dem Generationen von russischen und (nicht nur ost-)deutschen Naturwissenschaftlern und Ingenieuren ausgebildet wurden; enthält in vier Bänden (von denen die letzten beiden noch in Halbbände unterteilt sind) den gesamten klassischen Stoff der *Höheren Mathematik* und ist auch heute noch sehr gut zu lesen. Die HM I behandelt Stoff aus den Bänden I, II und III/1.

[R] R. ROTHE: *Höhere Mathematik,* Teubner

Sieben (dünne) Bände, zu denen allerdings auch Aufgaben- und Formelsammlung gehören, so daß die Darstellung insgesamt eher knapp ist. Für diese Vorlesung relevant sind die Bände II und III.

[Du] A. DUSCHEK: *Höhere Mathematik,* Springer Wien

Die österreichische Variante, vier recht ausführliche Bände, von denen hier vor allem die ersten beiden von Interesse sind.

[A] G. AUMANN: *Höhere Mathematik,* Bibliographisches Institut Mannheim

Drei relativ dünne Taschenbücher, deren erste beide trotzdem fast alles enthalten, was wir in dieser Vorlesung brauchen. Die Darstellung ist natürlich weniger ausführlich als in den dickleibigen Werken, aber das wird nicht jeder Student als Nachteil empfinden.

[C] R. COURANT: *Vorlesungen über Differential- und Integralrechnung I+II,* Springer

Ein Klassiker eines berühmten Mathematikers, auch heute noch sehr lesenswert. Trotz des Titels beschränkt sich das Buch nicht auf Analysis, sondern behandelt auch beispielsweise die Lineare Algebra in einem Umfang, der für diese Vorlesung völlig ausreicht.

Kapitel 1
Lineare Algebra

In der linearen Algebra geht es um Vektoren, Matrizen, Systeme linearer Gleichungen und den darauf aufbauenden Abstraktionen wie Vektorräumen und linearen Abbildungen. Genau wie Vektoren und Matrizen sind diese Methoden und Begriffe der linearen Algebra nicht wegzudenken aus der mehrdimensionalen Analysis, der Theorie der Differentialgleichungen sowie zahlreichen weiteren Gebieten innerhalb und außerhalb der Mathematik. Wir beginnen mit den wohl allgemein vertrauten

§1: Vektoren im $\mathbb{R}^3$

Unter einem *Pfeil* im $\mathbb{R}^3$ wollen wir die gerichtete Verbindungsstrecke zwischen zwei Punkten verstehen; er ist bestimmt nicht nur durch die Angabe von Anfangs- und Endpunkt, sondern beispielsweise auch durch seinen Anfangspunkt, seine Länge und seine Richtung, wobei diese Richtung allerdings für Pfeile der Länge Null nicht definiert ist.

Zwei Pfeile lassen sich addieren, falls der Endpunkt des ersten gleich dem Anfangspunkt des zweiten ist; die Summe ist dann derjenige Pfeil, der den Anfangspunkt des ersten mit dem Endpunkt des zweiten Pfeils verbindet. Auch läßt sich ein Pfeil mit einer reellen Zahl multiplizieren, wenn wir vereinbaren, daß das Ergebnis jeder Pfeil sein soll, der denselben Anfangspunkt und diesselbe Richtung hat wie der ursprüngliche Pfeil, dessen Länge aber mit der reellen Zahl multipliziert wurde. (Eine Multiplikation mit einer negativen Zahl soll dabei bedeuten, daß der Pfeil an seinem Anfangspunkt gespiegelt und dann mit dem Betrag der Zahl multipliziert wird.)

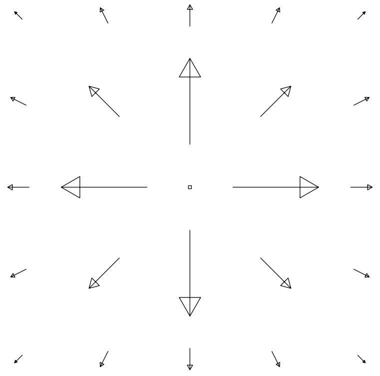


Abb. 1: Abstoßendes Feld einer Punktladung

Pfeile dieser Art können beispielsweise verwendet werden, um wie in Abbildung eins das Feld einer Punktladung zu veranschaulichen oder ein Kraftfeld *usw.* Trotzdem eignen sie sich nicht besonders gut als Grundbegriff einer mathematischen Theorie, da sich zwei Pfeile nur addieren lassen, wenn der Endpunkt des ersten gleich dem Anfangspunkt des zweiten ist, was etwa im Fall eines Kraft- oder elektromagnetischen Felds nicht auf den richtigen Additionsbegriff führt: Bei der Überlagerung der Felder zweier Punktladungen entsteht das Gesamtfeld schließlich daraus, daß man in jedem Punkt die beiden dort beginnenden Pfeile betrachtet und diese gemäß dem „Parallelogramm der Kräfte“ kombiniert, d.h. der eine Pfeil wird so lange verschoben, bis sein Anfangspunkt gleich dem Endpunkt des anderen Pfeils ist.

Wir definieren daher als neuen Begriff den des *Vektors,* wobei ein Vektor zwar wie ein Pfeil eine Länge und eine Richtung haben soll, aber keinen Anfangspunkt. Mathematisch exakt ausgedrückt ist also ein Vektor eine Äquivalenzklassen von Pfeilen, wobei zwei Pfeile genau dann äquivalent heißen, wenn sie dieselbe Länge und (so die Länge von Null verschieden ist) dieselbe Richtung haben.

Vektoren werden meist durch Fraktur- oder Fettbuchstaben bezeichnet; da sich Fettdruck schlecht auf einer Tafel realisieren läßt und die Hörer

dieser Vorlesung keine Frakturbuchstaben lernen wollen, verwenden wir hier stattdessen lateinische Buchstaben, die mit einem Pfeil überstrichen sind, also $\vec{u}, \vec{v}, \vec{w}, \ldots$. Die Länge eines Vektors $\vec{v}$ bezeichnen wir mit $|\vec{v}|$; man sollte sich kurz überlegen, daß das für „Vektoren“ in $\mathbb{R}^1$ genau mit dem gewohnten Betrag übereinstimmt.

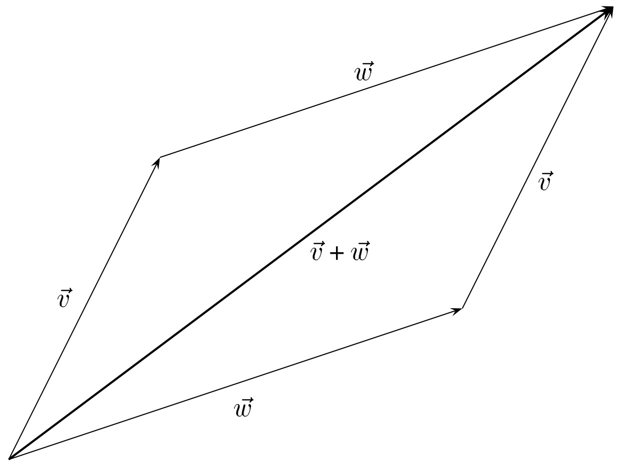


Abb. 2: Das „Parallelogramm der Kräfte“

Zwei Vektoren werden addiert, indem man sich zu jedem der beiden einen Pfeil wählt, wobei der Anfangspunkt des zweiten Pfeils gleich dem Endpunkt des ersten ist; die Summe dieser beiden Pfeile soll dann ein Pfeil für den Summenvektor sein. (Man veranschaulicht sich leicht, daß der resultierende Vektor nicht von der speziellen Wahl der Pfeile abhängt.) Aus dem „Parallelogramm der Kräfte“ in Abbildung zwei, das diese Addition veranschaulicht, liest man sofort ab, daß

$$\vec{v}+\vec{w}=\vec{w}+\vec{v}$$

für alle Vektoren $\vec{v}, \vec{w} \in \mathbb{R}^3$.

Genauso definiert man die Multiplikation eines Vektors mit einer reellen Zahl dadurch, daß man einen der Pfeile zum gegebenen Vektor mit der

Zahl multipliziert und den Vektor zum resultierenden Pfeil nimmt; die Rechenregel

$$\lambda(\vec{v}+\vec{w}) = \lambda\vec{v}+\lambda\vec{w}$$

folgt, indem man den Strahlensatz auf das Dreieck in Abbildung drei anwendet. Dabei sollen die fett eingezeichneten Vektoren die mit λ multiplizierten sein; im Bild ist $\lambda = 0{,}4$.

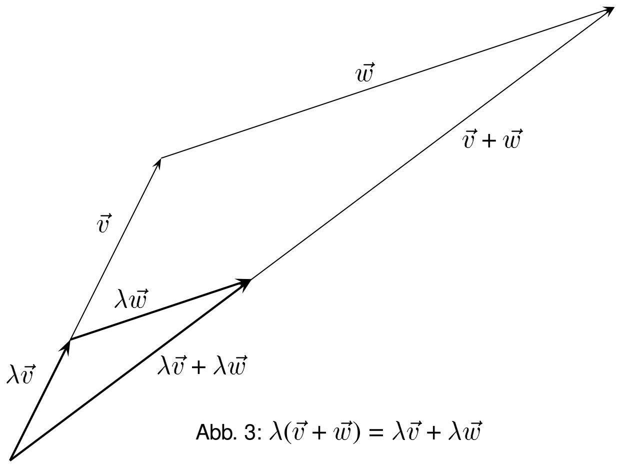


Abb. 3: $\lambda(\vec{v}+\vec{w}) = \lambda\vec{v}+\lambda\vec{w}$

Schließlich gibt es noch zwei *Produkte* von Vektoren im $\mathbb{R}^3$, das *Skalarprodukt* und das *Vektorprodukt.*

Beginnen wir mit dem Skalarprodukt. Es hat seinen Namen daher, daß es zwei Vektoren zu einem *Skalar,* d.h. einer reellen Zahl, verknüpft; diese wird mit $\vec{v}\cdot\vec{w}$ oder auch kurz $\vec{v}\vec{w}$ bezeichnet. Nach Definition ist

$$\vec{v}\cdot\vec{w} \underset{\text{def}}{=} |\vec{v}|\,|\vec{w}|\cos\measuredangle(\vec{v},\vec{w})\,,$$

wobei $\measuredangle(\vec{v},\vec{w})$ für den Winkel zwischen $\vec{v}$ und $\vec{w}$ steht. Da

$$\cos(2\pi-\varphi) = \cos(-\varphi) = \cos\varphi$$

ist, folgt sofort das *Kommutativgesetz*

$$\vec{v}\cdot\vec{w} = \vec{w}\cdot\vec{v}$$

für das Skalarprodukt; außerdem folgt wegen $\cos 0 = 1$, daß

$$\vec{v}\cdot\vec{v} = |\vec{v}|^2 \quad \text{oder} \quad |\vec{v}| = \sqrt{\vec{v}\cdot\vec{v}}$$

ist. Weiter ist $\cos(\pi/2) = \cos(3\pi/2) = 0$ und damit $\vec{v}\cdot\vec{w} = 0$, falls $\vec{v}$ und $\vec{w}$ aufeinander senkrecht stehen; falls keiner der beiden Vektoren der Nullvektor ist, stehen $\vec{v}$ und $\vec{w}$ *genau dann* senkrecht aufeinander, wenn $\vec{v}\cdot\vec{w} = 0$ ist.

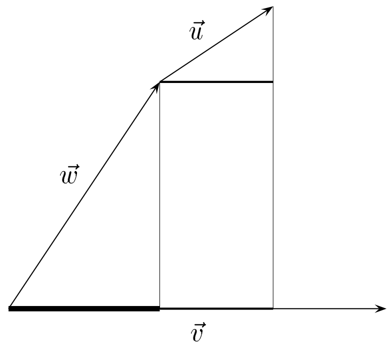


Abb. 4: Geometrische Interpretation des Skalarprodukts

Das Skalarprodukt kann auch geometrisch interpretiert werden: Da der Cosinus eines Winkels gleich Ankathete durch Hypothenuse ist, zeigt Abbildung vier, daß $|\vec{w}|\cos\measuredangle(\vec{v},\vec{w})$ gerade die Länge des senkrecht auf die von $\vec{v}$ erzeugte Gerade projizierten Vektors $\vec{w}$ ist; in der Zeichnung ist dies die fett eingezeichnete Strecke.

Addieren wir zu $\vec{w}$ einen weiteren Vektor $\vec{u}$, so ist auch hier wieder $|\vec{u}|\cos\measuredangle(\vec{v},\vec{u})$ die (in Abbildung vier halbfett eingezeichnete) Länge des auf die von $\vec{v}$ erzeugte Gerade projizierten Vektors $\vec{u}$ und entsprechendes gilt für $\vec{w}+\vec{u}$; wir erhalten somit die Regel

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

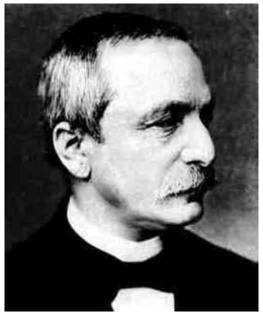
und damit wegen des Kommutativgesetzes auch

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

Zum Berechnen des Skalarprodukts ist die bisherige Definition für viele Fälle recht unhandlich, da Vektoren oft in einer solchen Weise gegeben sind, daß man den Winkel zwischen ihnen *nicht* ohne weiteres kennt. Typischerweise sind Vektoren bezüglich eines *Koordinatensystems* gegeben, etwa eines kartesischen Koordinatensystems aus drei paarweise aufeinander senkrecht stehenden Vektoren $\vec{e}_1, \vec{e}_2, \vec{e}_3$ der Länge eins. Für diese Vektoren läßt sich das Skalarprodukt leicht ausrechnen: Da zwei verschiedene stets aufeinander senkrecht stehen und jeder einzelne die Länge eins hat, ist

$$\vec{e}_i \cdot \vec{e}_j = \begin{cases} 0 & \text{falls } i \neq j \\ 1 & \text{falls } i = j \end{cases} .$$

Die rechte Seite wird im allgemeinen als KRONECKER-δ bezeichnet und abgekürzt durch δ_{ij}.



LEOPOLD KRONECKER (1823–1891) ist heute zwar vielen nur im Zusammenhang mit dem KRONECKER-δ bekannt, er war jedoch einer der bedeutendsten deutschen Mathematiker seiner Zeit. Seine Arbeiten befaßten sich mit Algebra, Zahlentheorie und Analysis, wobei er insbesondere die Verbindungen zwischen der Analysis und den beiden anderen Gebieten erforschte. Bekannt ist auch seine Ablehnung jeglicher mathematischer Methoden, die, wie die Mengenlehre oder Teile der Analysis, unendliche Konstruktionen verwenden. Er war deshalb mit vielen anderen bedeutenden Mathematikern seiner Zeit wie etwa CANTOR und WEIERSTRASS verfeindet.

Bezüglich eines solchen Koordinatensystems läßt sich ein beliebiger Vektor $\vec{v}$ in der Form

$$\vec{v} = v_1\vec{e}_1 + v_2\vec{e}_2 + v_3\vec{e}_3 \quad \text{mit} \quad v_i \in \mathbb{R}$$

schreiben; falls das Koordinatensystem feststeht (was meistens der Fall ist, da es sonst nicht viel nützt), schreiben wir dies auch kurz als

$$\vec{v} = \begin{pmatrix} v_1 \\ v_2 \\ b_3 \end{pmatrix} .$$

Für zwei Vektoren

$$\vec{v} = \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} \quad \text{und} \quad \vec{w} = \begin{pmatrix} w_1 \\ w_2 \\ w_3 \end{pmatrix}$$

ist dann

$$\vec{v} \cdot \vec{w} = (v_1\vec{e}_1 + v_2\vec{e}_2 + v_3\vec{e}_3) \cdot (w_1\vec{e}_1 + w_2\vec{e}_2 + w_3\vec{e}_3) ,$$

und nach obigen Rechenregeln läßt sich dies berechnen als

$$\vec{v} \cdot \vec{w} = \sum_{i=1}^{3} \sum_{j=1}^{3} v_i w_j \, \vec{e}_i \vec{e}_j = \sum_{i=1}^{3} \sum_{j=1}^{3} v_i w_j \delta_{ij} = \sum_{i=1}^{3} v_i w_i .$$

In dieser Form werden Skalarprodukte meistens ausgerechnet.

Das zweite Produkt von Vektoren im $\mathbb{R}^3$ ist das *Vektorprodukt* oder *Kreuzprodukt*. Wie die beiden Namen sagen, liefert es als Ergebnis einen Vektor und wird mit $\vec{v} \times \vec{w}$ bezeichnet. Dieser Vektor soll die Länge

$$|\vec{v} \times \vec{w}| = |\vec{v}| \, |\vec{w}| \, |\sin \measuredangle(\vec{v}, \vec{w})|$$

haben und senkrecht stehen sowohl auf $\vec{v}$ als auch auf $\vec{w}$. Insbesondere ist damit $\vec{v} \times \vec{w} = \vec{0}$ falls $\vec{v}$ und $\vec{w}$ dieselbe (oder entgegengesetzte) Richtung haben oder (mindestens) einer der beiden Vektoren der Nullvektor ist; in allen anderen Fällen spannen $\vec{v}$ und $\vec{w}$ eine Ebene auf, und $\vec{v} \times \vec{w}$ muß auf dieser Ebenen senkrecht stehen. Dazu gibt es zwei Möglichkeiten: Entweder er steht ab nach oben oder nach unten (wie immer man auch oben und unten festlegen möchte; für eine beliebige Ebene ist das natürlich völlig willkürlich).

Um diese beiden Fälle voneinander zu unterscheiden, benötigen wir einen neuen Begriff: Wir sagen, daß drei Vektoren $\vec{v}_1, \vec{v}_2$ und $\vec{v}_3$, die nicht in einer Ebenen liegen, ein *Rechtssystem* bilden, wenn sich die Finger der *rechten* Hand so ausrichten lassen, daß der Daumen in Richtung von $\vec{v}_1$ zeigt, der Zeigefinger in Richtung von $\vec{v}_2$ und der Mittelfinger in Richtung von $\vec{v}_3$. Alternativ kann man es, wenn der dritte Vektor so wie hier senkrecht auf den beiden ersten steht, auch so definieren, daß sich ein von $\vec{v}_1$ nach $\vec{v}_2$ gedrehter Korkenzieher in Richtung $\vec{v}_3$ in den Kork bohrt.

Ähnlich kann man das Rechtssystem auch mit Schrauben definieren; da es allerdings neben den (üblichen) Rechtsschrauben auch die (seltenen) Linksschrauben gibt, ist diese Definition eventuell zirkulär: Alles hängt davon ab, wie man Rechtsschrauben definiert.

Aus jeder dieser Regeln folgt sofort die *Antikommutativität* des Vektorprodukts:
$$\vec{v} \times \vec{w} = -\vec{w} \times \vec{v}\,.$$

Auch für das Vektorprodukt lassen sich leicht weitere Rechenregeln geometrisch ableiten: Da der Sinus eines Winkels gleich Gegenkathete durch Hypothenuse ist, ist in der von $\vec{v}$ und $\vec{w}$ aufgespannten Ebenen $|\vec{w}|\,|\sin\measuredangle(\vec{v},\vec{w})|$ gleich der Länge des auf die senkrecht auf $\vec{v}$ stehenden Geraden projizierten Vektors $\vec{w}$, das heißt also gleich der Höhe des in Abbildung fünf eingezeichneten Rechtecks. Die Länge des Vektors $\vec{v}\times\vec{w}$ ist daher gleich dem Flächeninhalt dieses Rechtecks und damit – wie eine Scherung zeigt – gleich der Fläche des von $\vec{v}$ und $\vec{w}$ aufgespannten Parallelogramms.

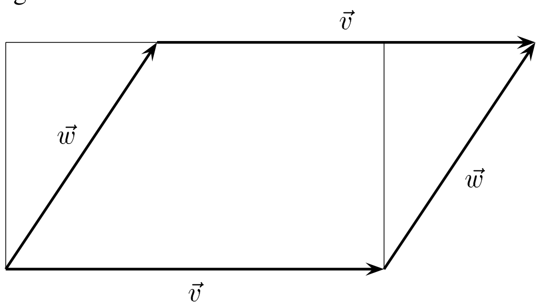


Abb. 5: Geometrische Interpretation des Vektorprodukts

Wie oben beim Skalarprodukt folgt daher, daß
$$\vec{v} \times (\vec{w} + \vec{u}) = \vec{v} \times \vec{w} + \vec{v} \times \vec{u}$$
und wegen der Antikommutativität entsprechend auch
$$(\vec{u} + \vec{v}) \times \vec{w} = \vec{u} \times \vec{w} + \vec{v} \times \vec{w}\,.$$

Um das Vektorprodukt in Koordinaten ausrechnen zu können, müssen wir zunächst die Produkte der $\vec{e}_i$ kennen; diese können wir aber nur ausrechnen, wenn wir wissen, ob $\vec{e}_1, \vec{e}_2$ und $\vec{e}_3$ ein Rechtssystem oder ein (analog zu definierendes) Linkssystem bilden. Wir nehmen an, daß sie ein Rechtssystem bilden; dann folgt sofort, daß $\vec{e}_1 \times \vec{e}_2 = \vec{e}_3$ ist, und nach einigen Fingerübungen auch, daß
$$\vec{e}_2 \times \vec{e}_3 = \vec{e}_1 \quad \text{und} \quad \vec{e}_1 \times \vec{e}_3 = -\vec{e}_2$$
ist. Die Produkte mit vertauschten Faktoren sind natürlich gerade das negative davon, und $\vec{e}_i \times \vec{e}_i = 0$. Für
$$\vec{v} = \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} \quad \text{und} \quad \vec{w} = \begin{pmatrix} w_1 \\ w_2 \\ w_3 \end{pmatrix}$$
ist also
$$\vec{v} \times \vec{w} = (v_1\vec{e}_1 + v_2\vec{e}_2 + v_3\vec{e}_3) \times (w_1\vec{e}_1 + w_2\vec{e}_2 + w_3\vec{e}_3)\,,$$
und nach obigen Rechenregeln ist dies gleich
$$\begin{aligned} &\sum_{i=1}^{3}\sum_{j=1}^{3} v_i w_j\, \vec{e}_i \times \vec{e}_j \\ &= (v_2w_3 - v_3w_2)\vec{e}_1 + (v_3w_1 - v_1w_3)\vec{e}_2 + (v_1w_2 - v_2w_1)\vec{e}_3\,. \end{aligned}$$
Dies läßt sich dadurch merken, daß man im Schema
$$\begin{array}{ccccccccc} \vec{e}_1 & & \vec{e}_2 & & \vec{e}_3 & & \vec{e}_1 & & \vec{e}_2 \\ & \searrow & & \times & & \times & & \swarrow & \\ v_1 & & v_2 & & v_3 & & v_1 & & v_2 \\ & \swarrow & & \times & & \times & & \searrow & \\ w_1 & & w_2 & & w_3 & & w_1 & & w_2 \end{array}$$
von $\vec{e}_i$ ausgeht und als dessen Koeffizient das Zweierprodukt entlang der schrägen Linie nach rechts unten *positiv* und das entlang der schrägen Linie nach links unten *negativ* nimmt.

Die Interpretation der Länge des Vektorprodukts als Fläche eines Parallelogramms wirft die Frage auf, ob man vielleicht auch Volumina auf diese Weise berechnen kann: Schließlich befinden wir uns im $\mathbb{R}^3$, wo es in erster Linie um Volumina geht und nicht um Flächeninhalte.

Nun liegen aber *zwei* Vektoren stets in einer Ebenen, können also keinen Körper definieren, der ein Volumen einschließen. Dies wird erst möglich, wenn noch ein *dritter* Vektor hinzukommt; dann entsteht ein dreidimensionales Analogon zum Parallelogramm, das in Abbildung fünf gezeigte *Parallelepiped,* das man sich als einen verzerrten Quader vorstellen kann, dessen rechte Winkel zu Winkeln einer beliebigen Größe wurden. Es wird auch als *Spat* bezeichnet, da Kalkspat und einige Feldspäte gelegentlich Kristalle dieser Form bilden. Aus dem Physikunterricht der Schule ist vielleicht der sogenannte Isländischen Doppelspat bekannt, eine der vielen Formen des Kalkspats: Dieser bildet meist sehr schöne Parallelepipede und ist vor allem bekannt durch seine ungewöhnlich stark ausgeprägte Doppelbrechung, die deshalb sehr häufig anhand dieser Kristalle demonstriert wird. (Sprachlich ist die Bezeichnung *Spat* für das Parallelepiped nicht sonderlich sinnvoll, denn die Endung „spat" beim Feldspat oder Kalkspat deutet nur auf die leichte Spaltbarkeit des Materials hin; sie hat nichts mit der geometrischen Form zu tun.)

Das Volumen eines Parallelepipeds läßt sich auf ähnliche Weise berechnen wie das eines Parallelogramms: Durch Scherung wie in Abbildung sieben kann man erreichen, daß der dritte Vektor senkrecht auf den beiden anderen steht; das Volumen des Parallelepipeds ist also gleich dem Volumen des entsprechenden Prismas. Das Volumen eines Prismas wiederum ist gleich der Grundfläche mal der Höhe, und die Grundfläche ist eine uns bekannte Parallelogrammfläche.

Bezeichnen wir also die Vektoren der Reihe nach als $\vec{u}, \vec{v}$ und $\vec{w}$, so ist die Fläche des von $\vec{u}$ und $\vec{v}$ aufgespannten Parallelogramms gleich $|\vec{u} \times \vec{v}|$. Die Höhe des Prisma ist gleich der Länge des auf eine Achse senkrecht zum Parallelogramm projezierten Vektors $\vec{w}$.

Diese Achse wird beispielsweise aufgespannt durch den Vektor $\vec{u} \times \vec{v}$; wir können die Höhe also berechnen als Betrag des Skalarprodukt von $\vec{w}$ mit einem Einheitsvektor, der dieselbe Richtung hat wie $\vec{u} \times \vec{v}$:

$$h = \left| \frac{\vec{u} \times \vec{v}}{|\vec{u} \times \vec{v}|} \cdot \vec{w} \right| = \frac{|(\vec{u} \times \vec{v}) \cdot \vec{w}|}{|\vec{u} \times \vec{v}|} .$$

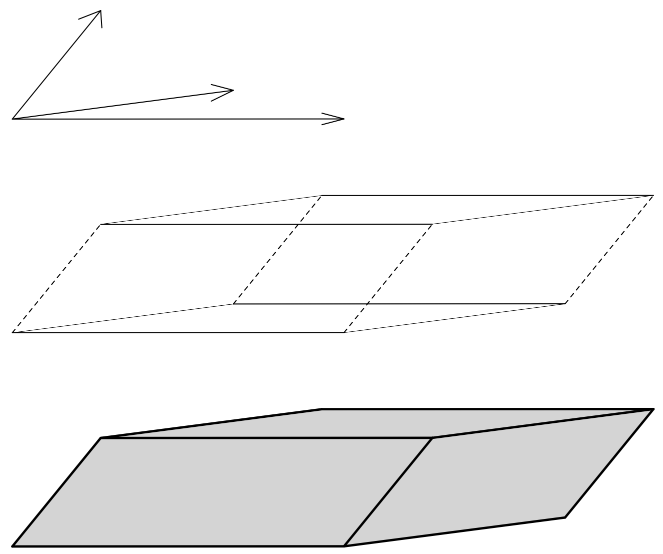


Abb. 6: Das von drei Vektoren aufgespannte Parallelepiped

Das Volumen des Prismas und damit des Parallelepipeds ist somit gleich

$$|\vec{u} \times \vec{v}| \cdot \frac{|(\vec{u} \times \vec{v}) \cdot \vec{w}|}{|\vec{u} \times \vec{v}|} = |(\vec{u} \times \vec{v}) \cdot \vec{w}| \ .$$

Wir bezeichnen deshalb das Produkt

$$(\vec{u} \times \vec{v}) \cdot \vec{w}$$

der drei Vektoren $\vec{u}, \vec{v}$ und $\vec{w}$ als das *Spatprodukt* dieser Vektoren; sein Betrag ist gleich dem Volumen des von den drei Vektoren aufgespannten Parallelepipeds. Da dieses Volumen nicht von der Reihenfolge der Vektoren abhängt, ist auch das Spatprodukt bis auf Vorzeichen unabhängig von der Reihenfolge der Vektoren; nachrechnen zeigt, daß

$$(\vec{u} \times \vec{v}) \cdot \vec{w} = (\vec{v} \times \vec{w}) \cdot \vec{u} = (\vec{w} \times \vec{u}) \cdot \vec{v}$$

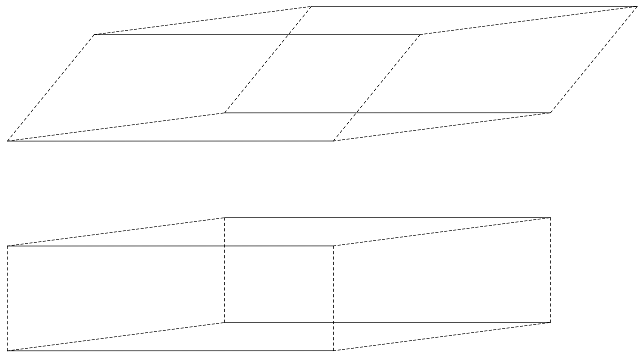


Abb. 7: Scherung eines Parallelepipeds zum Prisma

dasselbe Vorzeichen haben; das andere Vorzeichen haben die drei Produkte

$$(\vec v \times \vec u) \cdot \vec w = (\vec w \times \vec v) \cdot \vec u = (\vec u \times \vec w) \cdot \vec v \,,$$

was man schon daran sieht, daß hier die Reihenfolge der Faktoren des Kreuzprodukts umgekehrt wurde.

§2: Vektoren und Matrizen über beliebigen Körpern

Vektoren gibt es natürlich nicht nur im $\mathbb{R}^3$; man kann sich auch auf die ebenen Vektoren aus $\mathbb{R}^2$ beschränken, oder man kann – was in der Elektrotechnik wie auch in der Optik oft nützlich sind – auch Vektoren mit komplexen Zahlen als Einträgen betrachten. In der Informationstechnik schließlich spielen auch Vektoren eine Rolle, die nur Nullen und Einsen als Komponenten haben, Bitvektoren also. Der algebraische Begriff, unter dem wir die Skalarbereiche $\mathbb{R}, \mathbb{C}, \{0, 1\}$ und viele andere zusammenfassen können, ist der aus der Analysis I bekannte *Körper.*

Zur Erinnerung: Ein Körper ist eine Menge, für die zwei innere Verknüpfungen “+” und “·” definiert sind, die sich so verhalten, wie wir es von den reellen Zahlen her gewohnt sind. Formal heißt das

Definition: : Ein Körper k ist eine Menge zusammen mit zwei Abbildungen

$$+\colon k \times k \to k \quad \text{und} \quad \cdot\colon k \times k \to k \,,$$

genannt *Addition* und *Multiplikation,* für die gilt:

1.) Die Assoziativgesetze

$$(a+b)+c = a+(b+c) \quad \text{und} \quad (a\cdot b)\cdot c = a\cdot(b\cdot c) \qquad \text{für alle } a, b, c \in k$$

2. Es gibt zwei *verschiedene* Elemente $0, 1 \in k$, für die gilt

$$a + 0 = 0 + a = a \quad \text{und} \quad a \cdot 1 = 1 \cdot a = a \qquad \text{für alle } a \in k$$

3.) Zu jedem Element $a \in k$ gibt es ein Element $a' \in k$, so daß gilt

$$a + a' = a' + a = 0 \,.$$

Ist $a \neq 0$, so gibt es außerdem noch ein $a'' \in k$, so daß gilt

$$a \cdot a'' = a'' \cdot a = 1 \,.$$

4.) Die Kommutativgesetze

$$a + b = b + a \quad \text{und} \quad a \cdot b = b \cdot a \qquad \text{für alle } a, b \in k$$

5.) Das Distributivgesetz

$$a \cdot (b + c) = a \cdot b + a \cdot c \qquad \text{für alle } a, b, c \in k$$

Das Element a' aus 3.) wird üblicherweise als $-a$ bezeichnet und a'' als a^{-1}. Statt $a + (-b)$ schreibt man kurz $a - b$, und statt $a \cdot b^{-1}$ kurz a/b.

Überlegen wir uns ganz kurz, daß sich aus 1.) – 5.) in der Tat die üblichen Rechenregeln ableiten lassen: Zum Beispiel gilt

$$a + b = a + c \Longrightarrow b = c \,,$$

denn zunächst können wir auf beiden Seiten $-a$ addieren und erhalten

$$(-a) + (a + b) = (-a) + (a + c) \,,$$

was nach dem Asssoziativgesetz auch als

$$(-a + a) + b = (-a + a) + c$$

geschrieben werden kann. Nach 3.) ist $-a + a = 0$, und nach 2.) ist $0 + b = b$ und $0 + c = c$. Also ist $b = c$.

Genauso zeigt man, daß

$$a \cdot b = a \cdot c \Longrightarrow b = c \quad \text{falls} \quad a \neq 0\,,$$

denn dann kann man exakt genauso argumentieren, wenn man beide Seiten mit a^{-1} multipliziert.

Als nächstes folgt

$$a \cdot 0 = 0 \cdot a = 0 \quad \text{für alle } a \in k\,,$$

denn nach dem Distributivgesetz ist

$$a \cdot 0 + a \cdot 1 = a \cdot (0 + 1) = a \cdot 1 = 0 + a \cdot 1\,,$$

und daraus folgt wegen der gerade bewiesenen Aussage, daß $a \cdot 0 = 0$ ist. $0 \cdot a = 0$ folgt daraus durch einfache Anwendung des Kommutativitätsgesetzes.

Schließlich ist auch

$$-a = (-1) \cdot a \quad \text{für alle } a \in k\,,$$

denn

$$0 = a \cdot 0 = a \cdot (1 - 1) = a \cdot 1 + a \cdot (-1) = a + (-1) \cdot a\,.$$

Das soll uns für den Augenblick genügen; der Leser kann sich aber leicht überzeugen, daß man durch Herumspielen mit den obigen Axiomen über kurz oder lang in der Tat alle gängigen Rechenregeln erhält.

Die für uns wichtigsten Beispiele von Körpern sind wie gesagt $k = \mathbb{R}$ und $k = \mathbb{C}$; da aber ein Computer in keinem dieser beiden Körper exakt rechnen kann, ist auch der Körper $k = \mathbb{Q}$ wichtig sowie einige seiner Erweiterungskörper; diese finden zum Beispiel Anwendung in der Computergraphik. Für die bit-orientierte Informationsverarbeitung schließlich ist auch $k = \mathbb{F}_2 = \{0, 1\}$ wichtig; wir wollen uns daher überlegen, daß auch diese Menge zu einem Körper gemacht werden kann.

Für die Addition gibt es nicht viele Möglichkeiten: Wir müssen eines der beiden Elemente zum Neutralelement machen, wofür wir natürlich

sinnvollerweise die Null wählen. Alsdann ist nach Definition der Eigenschaften eines Neutralelements

$$0 + 0 = 0 \quad \text{und} \quad 0 + 1 = 1 + 0 = 1\,;$$

die einzige noch unbekannte Summe ist also $1+1$. Wäre $1+1 = 1 = 1+0$, müßte, nach dem was wir uns oben überlegt haben, $1 = 0$ sein, was nach Eigenschaft 2) aus der Definition eines Körpers nicht sein darf. Also müssen wir festlegen, daß $1 + 1 = 0$ ist.

Bei der Multiplikation ist ebenfalls alles festgelegt: Wie wir gerade gesehen haben, ist für jedes Element a eines Körpers k stets

$$a \cdot 0 = 0 \cdot a = 0\,.$$

also ist

$$0 \cdot 0 = 0, \quad 0 \cdot 1 = 0 \quad \text{und} \quad 1 \cdot 0 = 0\,.$$

Auf Grund von Eigenschaft 2.) aus der Definition eines Körpers ist außerdem $1 \cdot 1 = 1$, die Verknüpfungstabellen sehen also folgendermaßen aus:

+	0	1
0	0	1
1	1	0

·	0	1
0	0	0
1	0	1

Ein Leser, der bereits über Kenntnisse der Logik und/oder der Schaltungstechnik verfügt, wird hier sicherlich bekanntes entdecken:

- Falls man 1 als *wahr* und 0 als *falsch* interpretiert, ist „·" das logische Und, während „+" das *exklusive* logische Oder ist. (Für Altphilologen ist dies das lateinische *aut* im Gegensatz zum *vel;* wer sich eher für Logik oder Schaltalgebra interessiert, sollte zumindest eine der (äquivalenten) Bezeichnungen XOR oder *Antivalenz* schon einmal gehört haben.)
- Falls man ganze Zahlen in Binärdarstellung addieren möchte, ist für jede einzelne Binärstelle $x \cdot y$ der Übertrag, während $x + y$ bis auf den Übertrag der vorherigen Stelle gleich der Binärstelle des Ergebnisses ist. Man bezeichnet daher eine Schaltung, die $x + y$ und

$x \cdot y$ berechnet auch als einen *Halbaddierer;* der Volladdierer, der ein Bit plus dem Übertrag des vorherigen Bits verarbeitet, besteht aus zwei Halbaddierern und einem Oder-Gatter.

So seltsam dieser Körper auf den ersten Blick auch aussehen mag, ist er also doch unbestreitbar nützlich für verschiedenste Anwendungen. In der Kodierungstheorie spielen auch einige seiner Erweiterungskörper eine wichtige Rolle; dazu werden wir später noch kommen.

Im Augenblick wollen wir uns zunächst überlegen, was Vektoren über einem Körper k sein sollen.

Am anschaulichsten ist alles, wenn man sich zumindest am Anfang unter k den Körper der reellen Zahlen vorstellt, allerdings müssen wir darauf achten, daß ab jetzt nur noch Operationen und Begriffe erlaubt sind, die für beliebige Körper definiert sind: Winkel und Winkelfunktionen beispielsweise sind nun genausowenig definiert wie Quadratwurzeln (die schon für $k = \mathbb{Q}$ im allgemeinen nicht existieren) und damit Längen von Vektoren.

Ein Vektor $\vec{v} \in k^n$ ist daher zunächst ganz formal ein n-Tupel, das wir in Spaltenform als

$$\vec{v} = \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}$$

schreiben. Vektoren werden komponentententenweise addiert und mit Skalaren, d.h. Elementen des Körpers k, multipliziert: Für $\lambda \in k$ ist

$$\begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} + \begin{pmatrix} w_1 \\ \vdots \\ w_n \end{pmatrix} = \begin{pmatrix} v_1 + w_1 \\ \vdots \\ v_n + w_n \end{pmatrix} \quad \text{und} \quad \lambda \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} = \begin{pmatrix} \lambda v_1 \\ \vdots \\ \lambda v_n \end{pmatrix} .$$

Auch ein Skalarprodukt kann über die naheliegende Verallgemeinerung der für den $\mathbb{R}^3$ berechneten Formel als

$$\begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} \cdot \begin{pmatrix} w_1 \\ \vdots \\ w_n \end{pmatrix} = v_1 w_1 + \cdots + v_n w_n$$

definiert werden. Das Vektorprodukt ist allerdings eine Ausnahmeerscheinung der Dimension drei, die vor allem im Reellen nützlich ist; es läßt sich natürlich über seine Koordinatendarstellung

$$\begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} \times \begin{pmatrix} w_1 \\ w_2 \\ w_3 \end{pmatrix} = \begin{pmatrix} v_2 w_3 - v_3 w_2 \\ v_3 w_1 - v_1 w_3 \\ v_1 w_2 - v_2 w_1 \end{pmatrix}$$

für jeden k^3 definieren, aber in anderen Dimensionen gibt es keine entsprechende Verknüpfung.

Unmittelbar aus den Definitionen folgt, daß für alle Vektoren $\vec{v}, \vec{w} \in k^n$ und alle $\lambda \in k$ gilt:

$$\begin{aligned} \vec{v} + \vec{w} &= \vec{w} + \vec{v} \\ \lambda(\vec{v} + \vec{w}) &= \lambda\vec{v} + \lambda\vec{w} \\ \vec{v} \cdot \vec{w} &= \vec{w} \cdot \vec{v} \qquad \text{und} \\ \lambda(\vec{v} \cdot \vec{w}) &= (\lambda\vec{v}) \cdot \vec{w} . \end{aligned}$$

In vielen Fällen wird es notwendig sein, Vektoren weiteren Operationen zu unterziehen, beispielsweise können sie um einen bestimmten Winkel gedreht werden oder (etwa in der Computergraphik) vom $\mathbb{R}^3$ in den $\mathbb{R}^2$ projeziert werden. Fast alle derartigen Operationen sind Spezialfälle von linearen Abbildungen:

Definition: Eine lineare Abbildung $\varphi\colon k^m \to k^n$ ist eine Abbildung, die für alle $\lambda, \mu \in k$ und alle Vektoren $\vec{v}, \vec{w} \in k^m$ die Gleichung

$$\varphi(\lambda\vec{v} + \mu\vec{w}) = \lambda\varphi(\vec{v}) + \mu\varphi(\vec{w})$$

erfüllt.

Diese Definition läßt sich natürlich sofort durch vollständige Induktion ausdehnen zu der Aussage, daß φ genau dann linear ist, wenn für beliebige Körperelemente λ_i und Vektoren $\vec{v}_i$ auch

$$\varphi(\lambda_1\vec{v}_1 + \cdots \lambda_r\vec{v}_r) = \lambda_1\varphi(\vec{v}_1) + \cdots \lambda_r\varphi(\vec{v}_r)$$

ist; diese Formel gilt auch für $r = 1$, denn

$$\varphi(\lambda_1\vec{v}_1) = \varphi(\lambda_1\vec{v}_1 + 0\,\vec{v}_1) = \lambda_1\varphi(\vec{v}_1) + 0\,\varphi(\vec{v}_1) = \lambda_1\varphi(\vec{v}_1) .$$

Aus der Schule sind lineare Abbildungen wohl bekannt als Abbildungen von $\mathbb{R}$ nach $\mathbb{R}$ der Form $\varphi(x) = ax + b$. So eine Abbildung ist nur dann auch im hier definierten Sinn linear, wenn $b = 0$ ist, denn es muß gelten

$$b = \varphi(0) = \varphi(0+0) = \varphi(0) + \varphi(0) = b + b\,.$$

Die Abbildung $x \mapsto ax$ ist stets linear im hier definierten Sinn, denn

$$a(\lambda_1 x_1 + \lambda_2 x_2) = a\lambda_1 x_1 + a\lambda_2 x_2$$

nach dem Distributivgesetz. Mehr noch: *Die sämtlichen linearen Abbildungen $\varphi\colon k \to k$ haben die Form $x \mapsto ax$* mit geeigneten $a \in k$: Mit $a = \varphi(1)$ ist nämlich für ein beliebiges $x \in k$

$$\varphi(x) = \varphi(x \cdot 1) = x\varphi(1) = ax\,.$$

Die oben definierten linearen Abbildungen sind also im einfachsten Spezialfall $m = n = 1$ gerade die Abbildungen $x \mapsto ax$ mit $a \in k$.

Wir wollen uns überlegen, daß man auch beliebige lineare Abbildungen $\varphi\colon k^m \to k^n$ in ähnlicher Form schreiben kann, wenn man das Körperelement a durch eine sogenannte Matrix A ersetzt und die Körpermultiplikation durch eine noch zu definierende Matrix-Vektor-Multiplikation.

Dazu definieren wir in k^m *Einheitsvektoren* $\vec{e}_i$ durch

$$\vec{e}_1 = \begin{pmatrix}1\\0\\0\\\vdots\\0\end{pmatrix}, \quad \vec{e}_2 = \begin{pmatrix}0\\1\\0\\\vdots\\0\end{pmatrix}, \quad \ldots, \quad \vec{e}_m = \begin{pmatrix}0\\0\\0\\\vdots\\1\end{pmatrix};$$

mit diesen läßt sich ein beliebiger Vektor $\vec{v} \in k^m$ schreiben als

$$\vec{v} = \begin{pmatrix}v_1\\\vdots\\v_n\end{pmatrix} = v_1\vec{e}_1 + \cdots + v_n\vec{e}_n\,.$$

Für eine lineare Abbildung $\varphi\colon k^m \to k^n$ ist daher

$$\varphi(\vec{v}) = v_1\varphi(\vec{e}_1) + \cdots + v_m\varphi(\vec{e}_m)\,.$$

Falls wir die Vektoren $\varphi(e_i)$ kennen, können wir also $\varphi(\vec{v})$ für jeden Vektor $\vec{v} \in k^m$ ausrechnen: Mit

$$\varphi(\vec{e}_i) = \begin{pmatrix}a_{1i}\\\vdots\\a_{ni}\end{pmatrix},$$

ist

$$\varphi(\vec{v}) = v_1\begin{pmatrix}a_{11}\\\vdots\\a_{n1}\end{pmatrix} + \cdots + v_m\begin{pmatrix}a_{1m}\\\vdots\\a_{nm}\end{pmatrix} = \begin{pmatrix}a_{11}v_1 + \cdots + a_{1m}v_m\\\vdots\\a_{n1}v_1 + \cdots + a_{nm}v_m\end{pmatrix}.$$

Die Elemente a_{ij} fassen wir zusammen zu einem zweidimensionalen Schema

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

der *Abbildungsmatrix* von φ. Ihre Spalten sind, wenn man sie als Vektoren auffaßt, gerade die Bilder der Einheitsvektoren.

Definition: Eine $n \times m$-Matrix A über dem Körper k ist eine zweidimensionale Anordnung von Körperelementen $a_{ik} \in k$ in n Zeilen und m Spalten, d.h.

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

Die Menge aller $n \times m$-Matrizen A über k bezeichnen wir mit $k^{n\times m}$.

Wie obige Rechnung zeigt, ist eine lineare Abbildung φ durch ihre Abbildungsmatrix eindeutig bestimmt. Um dies in möglichst kompakter Weise darzustellen, definieren wir das Produkt einer $n \times m$-Matrix $A \in k^{n\times m}$ mit einem Vektor $\vec{v} \in k^m$ als

$$A\vec{v} \underset{\text{def}}{=} \begin{pmatrix}a_{11}v_1 + \cdots + a_{1m}v_m\\\vdots\\a_{n1}v_1 + \cdots + a_{nm}v_m\end{pmatrix} = \begin{pmatrix}\sum_{j=1}^m a_{1j}v_j\\\vdots\\\sum_{j=1}^m a_{nj}v_j\end{pmatrix};$$

damit läßt sich also jede lineare Abbildung $\varphi\colon k^m \to k^n$ in der Form

$$\varphi(\vec{v}) = A\vec{v}$$

schreiben, wobei $A \in k^{n\times m}$ die Abbildungsmatrix von φ ist.

Ist umgekehrt $A \in k^{n\times m}$ eine beliebige Matrix, so überzeugt man sich leicht, daß die Abbildung

$$\varphi\colon \begin{cases} k^n \to k^m \\ \vec{v} \mapsto A\vec{v} \end{cases}$$

linear ist; die linearen Abbildungen $k^m \to k^n$ entsprechen also eineindeutig den Matrizen aus $k^{n\times m}$.

Genau wie Vektoren lassen sich auch $n \times m$-Matrizen addieren und mit Körperelementen $\lambda \in k$ multiplizieren gemäß

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

und

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

Auch hier folgen wieder unmittelbar die Rechenregeln

$$A + B = B + A \qquad \text{und}$$
$$\lambda(A + B) = \lambda A + \lambda B .$$

Wir hätten Addition von Matrizen und Multiplikation mit einem Skalar aber auch anders definieren können: Sind nämlich $\varphi, \psi\colon k^m \to k^n$ zwei lineare Abbildungen und ist $\lambda \in k$ ein Skalar, so sind auch

$$\varphi+\psi\colon \begin{cases} k^m \to k^n \\ \vec{v} \mapsto \varphi(\vec{v}) + \psi(\vec{v}) \end{cases} \quad \text{und} \quad \lambda\varphi\colon \begin{cases} k^m \to k^n \\ \vec{v} \mapsto \lambda\varphi(\vec{v}) \end{cases}$$

lineare Abbildungen, und ihre Abbildungsmatrizen sind offenbar gerade $A + B$ und λA, wobei A und B die Abbildungsmatrizen von φ und ψ bezeichnen.

Diese zweite Vorgehensweise führt uns in natürlicher Weise auf ein Produkt von Matrizen:

Zwei lineare Abbildungen $\varphi\colon k^m \to k^n$ und $\psi\colon k^p \to k^m$ lassen sich verknüpfen zu einer Abbildung

$$\varphi\circ\psi\colon \begin{cases} k^p \to k^n \\ \vec{v} \mapsto \varphi\bigl(\psi(\vec{v})\bigr) \end{cases} .$$

Die zu definierende Matrixmultiplikation soll dieser Hintereinanderausführung von Abbildungen entsprechen; ist also $A \in k^{n\times m}$ die Abbildungsmatrix zu φ und $B \in k^{m\times p}$ die zu ψ, so soll $AB \in k^{n\times p}$ die Abbildungsmatrix zu $\varphi\circ\psi$ sein.

Nach Definition ist die i-te Spalte der Abbildungsmatrix von $\varphi\circ\psi$ gleich dem Bild des i-ten Einheitsvektors $\vec{d}_i$ von k^p unter der Abbildug $\varphi\circ\psi$, also $\varphi\bigl(\psi(\vec{d}_i)\bigr)$. Dabei ist

$$\psi(\vec{d}_i) = \begin{pmatrix} b_{1i} \\ \vdots \\ b_{mi} \end{pmatrix} = b_{1i}\vec{e}_1 + \dots + b_{mi}\vec{e}_m$$

gleich der i-ten Spalte der Abbildungsmatrix B von ψ, und da der Einheitsvektor $\vec{e}_j$ von k^m von φ auf die j-te Spalte der Abbildungsmatrix A

von φ abgebildet wird, ist

$$\varphi\big(\psi(\vec{d}_i)\big) = b_{1i}\begin{pmatrix} a_{11} \\ \vdots \\ a_{n1}\end{pmatrix} + \cdots + b_{mi}\begin{pmatrix} a_{1m} \\ \vdots \\ a_{nm}\end{pmatrix} = \begin{pmatrix} a_{11}b_{1i} + \cdots + a_{1m}b_{mi} \\ \vdots \\ a_{n1}b_{1i} + \cdots + a_{nm}b_{mi}\end{pmatrix},$$

die Abbildungsmatrix C von $\varphi \circ \psi$ hat also an der Stelle hi den Eintrag

$$c_{hi} = a_{h1}b_{1i} + \cdots + a_{hm}b_{mi} = \sum_{j=1}^{m} a_{hj}b_{ji}\,.$$

Diese Abbildungsmatrix soll, wie gesagt, als Produkt von A und B definiert werden, wir *definieren* daher:

Definition: $A = (a_{i\ell}) \in k^{n\times m}$ und $B = (b_{\ell j}) \in k^{m\times p}$ seien zwei Matrizen derart, daß die Spaltenzahl m von A mit der Zeilenzahl von B übereinstimmt. Dann bezeichnen wir die Matrix $C = (c_{ij}) \in k^{n\times p}$ mit

$$c_{ij} = \sum_{\ell=1}^{m} a_{i\ell}b_{\ell j}$$

als *Produktmatrix* $C = AB$ von A und B.

Im Gegensatz zu den meisten bislang aufgetretenen Produkten ist dieses Matrixprodukt im allgemeinen *nicht* kommutativ: Falls nicht zufälligerweise $n = p$ sein sollte, ist das Matrixprodukt BA nicht einmal definiert, geschweige denn gleich AB. Allgemein ist Kommutativität bei der Hintereinanderausführung von Abbildungen eine sehr seltene Ausnahmeerscheinung; schließlich ist auch $\sin(x^2) \neq \sin^2 x$ für fast jedes x, und so ist auch bei Matrizen, selbst wenn beide Produkte definiert sind, im allgemeinen $AB \neq BA$. Beispielsweise ist

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

aber

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

Ansonsten gelten aber doch die meisten bekannten Rechenregeln, beispielsweise das *Assoziativitätsgesetz*

$$A(BC) = (AB)C \quad \text{für alle } A \in k^{n\times m},\ B \in k^{m\times p},\ C \in k^{p\times q}\,.$$

Es ist durchaus möglich (und verglichen mit manch anderen Dingen sogar nicht einmal so extrem aufwendig), dieses Gesetz nach obiger Formel explizit nachzurechnen. Bevor wir uns das antun, sollten wir uns aber daran erinnern, wo das Matrixprodukt eigentlich herkommt: Matrizen entsprechen umkehrbar eindeutig linearen Abbildungen, und das Matrixprodukt entspricht deren Hintereinanderausführung. Für die Hintereinanderausführung von Abbildungen (egal ob linear oder nicht) ist das Assoziativgesetz aber trivial: Sind etwa

$$\varphi\colon k^m \to k^n, \quad \psi\colon k^n \to k^p \quad \text{und} \quad \omega\colon k^q \to k^p$$

drei lineare Abbildungen mit Abbildungsmatrizen A, B, C, so ist für jeden Vektor $\vec{v} \in k^q$ sowohl

$$\big(\varphi\circ(\psi\circ\omega)\big)(\vec{v}) = \varphi\big(\psi\circ\omega)(\vec{v})\big) = \varphi\Big(\psi\big(\omega(\vec{v})\big)\Big)$$

als auch

$$\big((\varphi\circ\psi)\circ\omega\big)\big(\vec{v}\big) = (\varphi\circ\psi)\big(\omega(\vec{v})\big) = \varphi\Big(\psi\big(\omega(\vec{v})\big)\Big)\,,$$

d.h. für die Hintereinanderausführung von Abbildungen (egal ob linear oder nicht) ist das Assoziativgesetz

$$\varphi\circ(\psi\circ\omega) = (\varphi\circ\psi)\circ\omega$$

automatisch erfüllt.

Da nun $A(BC)$ die Abbildungsmatrix von $\varphi\circ(\psi\circ\omega)$ ist und $(AB)C$ die von $(\varphi\circ\psi)\circ\omega$ und da diese beiden Abbildungen übereinstimmen, müssen auch die Abbildungsmatrizen gleich sein, wir haben also gezeigt, daß

$$A(BC) = (AB)C \quad \text{für alle } A \in k^{n\times m},\ B \in k^{m\times p},\ C \in k^{p\times q}\,,$$

ohne daß wir ein einziges Matrixprodukt explizit ausrechnen mußten.

Entsprechend folgen auch die Rechenregeln

$$A(B_1 + B_2) = AB_1 + AB_2 \quad \text{und} \quad (A_1 + A_2)B = A_1B + A_2B$$

aus den entsprechenden Rechenregeln

$$\varphi\circ(\psi_1+\psi_2)=\varphi\circ\psi_1+\varphi\circ\psi_2 \quad\text{und}\quad (\varphi_1+\varphi_2)\circ\psi=\varphi_1\circ\psi+\varphi_2\circ\psi\,,$$

die sich wiederum leicht und ohne Rechnung überprüfen lassen:

$$\begin{aligned}\bigl(\varphi\circ(\psi_1+\psi_2)\bigr)(\vec v)&=\varphi\bigl(\psi_1(\vec v)+\psi_2(\vec v)\bigr)=\varphi\bigl(\psi_1(\vec v)\bigr)+\varphi\bigl(\psi_2(\vec v)\bigr)\\&=(\varphi\circ\psi_1)(\vec v)+(\varphi\circ\psi_2)(\vec v)=(\varphi\circ\psi_1+\varphi\circ\psi_2)(\vec v)\end{aligned}$$

und

$$\begin{aligned}\bigl((\varphi_1+\varphi_2)\circ\psi\bigr)(\vec v)&=(\varphi_1+\varphi_2)\bigl(\psi(\vec v)\bigr)=\varphi_1\bigl(\psi(\vec v)\bigr)+\varphi_2\bigl(\psi(\vec v)\bigr)\\&=(\varphi_1\circ\psi)(\vec v)+(\varphi_2\circ\psi)(\vec v)=(\varphi_1\circ\psi+\varphi_2\circ\psi)(\vec v)\,.\end{aligned}$$

Da in der obigen Formel für die Matrixmultiplikation alle $b_{\ell j}$ linear in den Ausdrücken für c_{ij} vorkommen *usw.,* hätten sich die beiden letzten Rechenregeln auch einfach direkt nachrechnen lassen.

Völlig problemlos sind Regeln wie

$$(\lambda A)B=\lambda(AB)=A(\lambda B)\quad\text{für alle }\lambda\in k\,,$$

und für die *Einheitsmatrix*

$$E=\begin{pmatrix}1&0&\dots&0\\0&1&\dots&0\\\vdots&\vdots&\ddots&\vdots\\0&0&\dots&1\end{pmatrix}\in k^{n\times n}$$

folgt ebenfalls sofort durch Nachrechnen wie auch durch Interpretation von E als Abbildungsmatrix der identischen Abbildung $k^n\to k^n$, die jeden Vektor auf sich selbst abbildet, daß

$$A\cdot E=A\qquad\text{für alle }A\in k^{m\times n}$$

und

$$E\cdot A=A\qquad\text{für alle }A\in k^{n\times m}$$

ist.

Zumindest für abstrakte Zusammenhänge kann man also das Rechnen mit der doch etwas unhandlichen Formel für das Matrixprodukt oft auch vermeiden; irgendwann gibt es dann aber doch keine Alternative mehr,

und man muß rechnen. Das Auswendiglernen der Regel für das Matrixprodukt ist recht fehleranfällig; für praktische Rechnungen empfiehlt es sich als Eselsbrücke den zweiten Faktor des Produkts höher zu stellen nach dem Schema

$$\begin{array}{cc} & \begin{pmatrix}\cdots & b_{1j} & \cdots\\ & \vdots & \\ \cdots & b_{mj} & \cdots\end{pmatrix}\\ \begin{pmatrix}\vdots & & \vdots\\ a_{i1} & \cdots & a_{im}\\ \vdots & & \vdots\end{pmatrix} & \begin{pmatrix} & \vdots & \\ \cdots & \sum\limits_{\ell=1}^{m} a_{i\ell}b_{\ell j} & \cdots\\ & \vdots & \end{pmatrix}\end{array}\quad;$$

dadurch behält man einerseits den Überblick, welcher Rechenschritt als nächster auszuführen ist; außerdem sieht man anhand dieser Darstellung, daß der Eintrag an der Stelle (i,j) der Produktmatrix gerade das Skalarprodukt des i-ten Zeilenvektors von A und mit dem j-ten Spaltenvektors von B ist:

$$c_{ij}=\begin{pmatrix}a_{i1}\\\vdots\\a_{im}\end{pmatrix}\cdot\begin{pmatrix}b_{1j}\\\vdots\\b_{mj}\end{pmatrix}\,.$$

Entsprechend steht in der j-ten Spalte der Produktmatrix AB gerade das Produkt von A mit der als Vektor aufgefaßten j-ten Spalte von B.

Dies zeigt insbesondere auch, daß man das oben definierte Produkt einer Matrix mit einem Vektor als Spezialfall des Matrixprodukts auffassen kann, weshalb Vektoren vor allem in der physikalischen und technischen Literatur gerne mit $n\times 1$-Matrizen identifiziert werden. Auch das Skalarprodukt zweier Vektoren läßt sich als Spezialfall des Matrixprodukts interpretieren, wenn man den ersten Vektor als eine $1\times n$- und den zweiten als eine $n\times 1$-Matrix betrachtet.

§3: Lineare Gleichungssysteme

Die häufigste und wohl auch wichtige Anwendung der linearen Algebra ist das Lösen linearer Gleichungssysteme. Unter einem linearen