Hier müssen wir vor der Anwendung des GAUSS-Algorithmus offensichtlich Zeilen vertauschen, z.B. die erste und die zweite. Dies entspricht einer Multiplikation mit

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

und führt auf

$$A' = PA = \begin{pmatrix} 3 & 4 & 5 \\ 0 & 1 & 2 \\ 6 & 7 & 8 \end{pmatrix} .$$

Wir schreiben die Einheitsmatrix daneben:

$$\begin{pmatrix} 3 & 4 & 5 & 1 & 0 & 0 \\ 0 & 1 & 2 & 0 & 1 & 0 \\ 6 & 7 & 8 & 0 & 0 & 1 \end{pmatrix} .$$

Die Sechs links unten wird eliminiert durch Subtraktion der zweifachen ersten Zeile von der letzten:

$$\begin{pmatrix} 3 & 4 & 5 & 1 & 0 & 0 \\ 0 & 1 & 2 & 0 & 1 & 0 \\ 0 & -1 & -2 & -2 & 0 & 1 \end{pmatrix}$$

Die Endgestalt entsteht daraus, wenn man nun noch die zweite Zeile zur dritten addiert:

$$\begin{pmatrix} 3 & 4 & 5 & 1 & 0 & 0 \\ 0 & 1 & 2 & 0 & 1 & 0 \\ 0 & 0 & 0 & -2 & 1 & 1 \end{pmatrix}$$

Hier steht links die Matrix R, rechts steht L, d.h.

$$R = \begin{pmatrix} 3 & 4 & 5 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{pmatrix} \quad \text{und} \quad L = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -2 & 1 & 1 \end{pmatrix} .$$

In der Tat rechnet man leicht nach, daß

$$LPA = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -2 & 1 & 1 \end{pmatrix} \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & 1 & 2 \\ 3 & 4 & 5 \\ 6 & 7 & 8 \end{pmatrix}$$

$$= \begin{pmatrix} 3 & 4 & 5 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{pmatrix} = R$$

ist.

§4: Determinanten

Determinanten liefern ein numerisches Kriterium für die lineare Abhängigkeit von n Vektoren im n-dimensionalen Raum. Zum besseren Verständnis betrachten wir zunächst den zwei- und den dreidimensionalen Fall.

a) Das Skalarprodukt und Vektorprodukt im $\mathbb{R}^3$

Aus der Schule bekannt sind zwei Produkte von Vektoren, die wir bei der Definition des Vektorraums nicht mitdefiniert haben: Das Skalarprodukt und das Vektorprodukt.

Beginnen wir mit dem Skalarprodukt. Es hat seinen Namen daher, daß es zwei Vektoren zu einem *Skalar* 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 Betragsstriche die Länge eines Vektors bezeichnen sollen und $\measuredangle(\vec{v}, \vec{w})$ 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.

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.

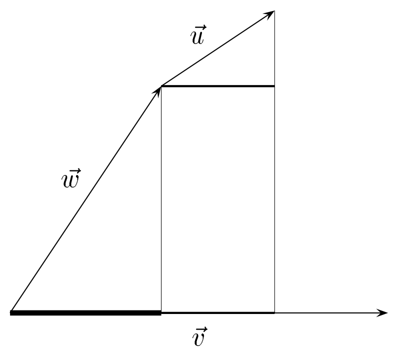

Abb. 11: Geometrische Interpretation des Skalarprodukts

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. Um gut rechnen zu können, wählen wir eine Basis 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

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

(δ_{ij} ist das aus §2 bekannte KRONECKER-δ)

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.

Im Dreidimensionalen (und nur dort) gibt es auch noch das sogenannte *Vektorprodukt* oder *Kreuzprodukt.* Wie die beiden Namen sagen, liefert es als Ergebnis einen Vektor und wird mit $\vec{v} \times \vec{w}$ bezeichnet.

Der Vektor $\vec{v} \times \vec{w}$ 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 $\vec{v} \times \vec{w} = \vec{0}$ genau dann, falls $\vec{v}$ und $\vec{w}$ dieselbe (oder entgegengesetzte) Richtung haben, oder (mindestens) einer der beiden Vektoren der Nullvektor ist, d.h. also genau dann, wenn $\vec{v}$ und $\vec{w}$ linear abhängig sind.

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:

$$\vec{v} \times \vec{w} = -\vec{w} \times \vec{v} \,.$$

Weitere Rechenregeln für das Vektorprodukt lassen sich leicht 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.

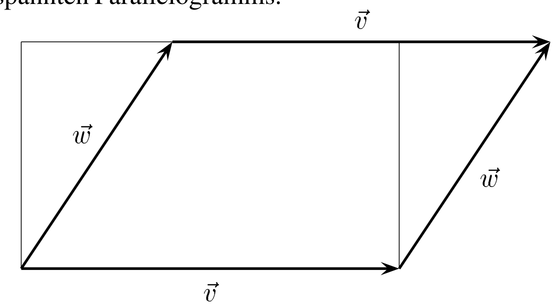

Abb. 12: Geometrische Interpretation des Vektorprodukts

Damit folgt, daß

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

und wegen der Antikommutativität entsprechend auch

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

ist.

Um das Vektorprodukt in Koordinaten ausrechnen zu können, müssen wir zunächst die Produkte der $\vec{e}_i$ kennen; diese können wir ausrechnen, sobald 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

$$\begin{aligned} &\sum_{i=1}^{3}\sum_{j=1}^{3} v_i w_j\, \vec{e}_i \times \vec{e}_j \\ &= (v_2w_3 - v_3w_2)\vec{e}_1 + (v_3w_1 - v_1w_3)\vec{e}_2 + (v_1w_2 - v_2w_1)\vec{e}_3\,, \end{aligned}$$

d.h.

$$\begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} \times \begin{pmatrix} w_1 \\ w_2 \\ w_3 \end{pmatrix} = \begin{pmatrix} v_2w_3 - v_3w_2 \\ v_3w_1 - v_1w_3 \\ v_1w_2 - v_2w_1 \end{pmatrix}\,.$$

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.

In diesem Paragraphen geht es uns um ein Kriterium, wann n Vektoren in k^n linear abhängig sind. Das Kreuzprodukt sagt uns, wann zwei Vektoren im $\mathbb{R}^3$ linear abhängig sind. Um ein Kriterium für zwei Vektoren $\vec{v} = \binom{v_1}{v_2}$ und $\vec{w} = \binom{w_1}{w_2}$ im $\mathbb{R}^2$ zu bekommen, betten wir diese ein in den $\mathbb{R}^3$, indem wir eine dritte Komponente Null hinzufügen; dann sind $\vec{v}$ und $\vec{w}$ genau dann linear abhängig, wenn

$$\begin{pmatrix} v_1 \\ v_2 \\ 0 \end{pmatrix} \times \begin{pmatrix} w_1 \\ w_2 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ v_1w_2 - v_2w_1 \end{pmatrix}$$

der Nullvektor ist, wenn also $v_1w_2 - v_2w_1 = 0$ ist.

Definition: Für zwei Vektoren $\vec{v} = \binom{v_1}{v_2}$ und $\vec{w} = \binom{w_1}{w_2}$ aus dem Vektorraum k^2 über einem Körper k bezeichnen wir die Zahl

$$\det(\vec{v}, \vec{w}) \underset{\text{def}}{=} \begin{vmatrix} v_1 & w_1 \\ v_2 & w_2 \end{vmatrix} \underset{\text{def}}{=} v_1w_2 - v_2w_1$$

als *Determinante* von $\vec{v}$ und $\vec{w}$.

Auch unabhängig von der Herleitung über das Kreuzprodukt ist klar, daß $\vec{v}$ und $\vec{w}$ genau dann linear abhängig sind, wenn $\det(\vec{v}, \vec{w})$ verschwindet. Genauer haben wir folgende drei Eigenschaften:

(D1) $\det(\vec{u}, \vec{v}) = -\det(\vec{v}, \vec{u})$

(D2) $\det(\vec{u}, \vec{v})$ ist linear in jedem ihrer beiden Argumente, d.h.

$$\det(\lambda\vec{u}_1 + \mu\vec{u}_2, \vec{v}) = \lambda \det(\vec{u}_1, \vec{v}) + \mu \det(\vec{u}_2, \vec{v})$$

und

$$\det(\lambda(\vec{u}, \lambda\vec{v}_1 + \mu\vec{v}_2) = \lambda \det(\vec{u}, \vec{v}_1) + \mu \det(\vec{u}, \vec{v}_2)\,.$$

(D3) $\det(\vec{u}, \vec{v}) = 0$ genau dann, wenn $\vec{u}$ und $\vec{v}$ linear abhängig sind.

b) Das Spatprodukt

Die Interpretation der Länge des Vektorprodukts als Fläche eines Parallelogramms wirft die Frage auf, ob man vielleicht auch Volumina auf diese Weise berechnen kann; dies würde dann auch auf ein Kriterium für die lineare Abhängigkeit liefern, denn drei Vektoren im $\mathbb{R}^3$ sind genau

dann linear unabhängig, wenn sie einen Körper mit Volumen ungleich null aufspannen.

Zwei Vektoren liegen stets in einer Ebenen; kommt noch ein *dritter* Vektor dazu, 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 14 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 $\vec{u} \times \vec{v}$:

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

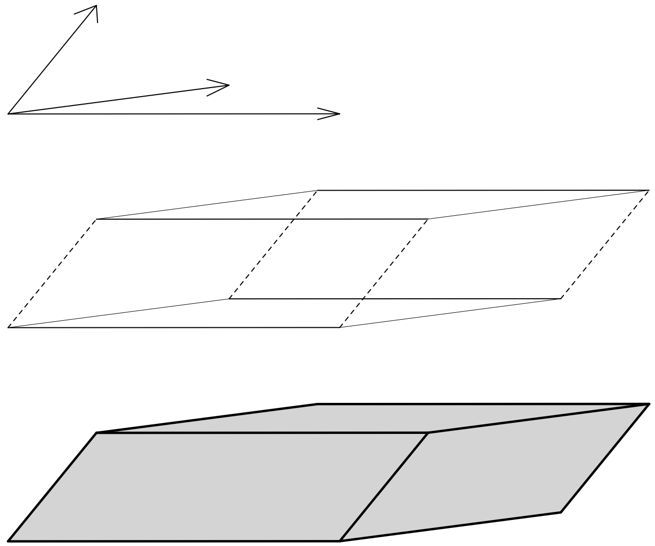

Abb. 13: Das von drei Vektoren aufgespannte Parallelepiped

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

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

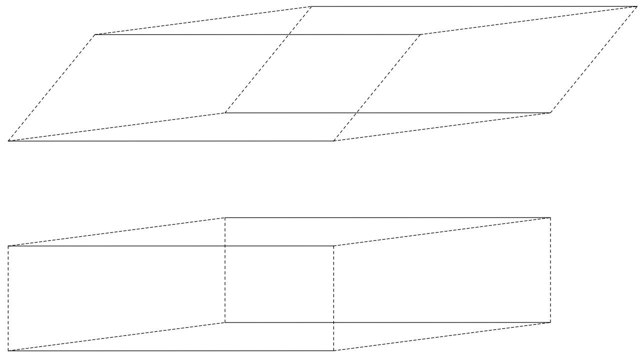

Abb. 14: Scherung eines Parallelepipeds zum Prisma

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.

In Koordinaten ausgedrückt ist $(\vec{u}\times\vec{v})\cdot\vec{w}$

$$=\left(\begin{pmatrix}u_1\\u_2\\u_3\end{pmatrix}\times\begin{pmatrix}v_1\\v_2\\v_3\end{pmatrix}\right)\cdot\begin{pmatrix}w_1\\w_2\\w_3\end{pmatrix}=\begin{pmatrix}u_2v_3-u_3v_2\\u_3v_1-u_1v_3\\u_1v_2-u_2v_1\end{pmatrix}\cdot\begin{pmatrix}w_1\\w_2\\w_3\end{pmatrix}$$

$$=u_2v_3w_1-u_3v_2w_1+u_3v_1w_2-u_1v_3w_2+u_1v_2w_e-u_2v_1w_3\ .$$

Diese Formel läßt sich leicht merken nach der folgenden Regel von Pierre Sarrus (1798–1861), von der wir bereits beim Kreuzprodukt einen Spezialfall kennengelernt haben: Schreibt man die Komponenten der Vektoren in der Form

$$\begin{array}{ccccccccc}
u_1 & & v_1 & & w_1 & & u_1 & & v_1\\
 & \searrow & & \times & & \times & & \swarrow & \\
u_2 & & v_2 & & w_2 & & u_2 & & v_2\\
 & \swarrow & & \times & & \times & & \searrow & \\
u_3 & & v_3 & & w_3 & & u_3 & & v_3
\end{array}\ ,$$

so sind die Produkte entlang der drei schrägen Linien von links oben nach rechts unten, der Hauptdiagonalen und ihrer Parallelen also, positiv zu rechnen und die entlang der drei schrägen Linien von rechts oben nach links unten negativ.

Definition: Für drei Vektoren

$$\vec{u} = \begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix}, \quad \vec{v} = \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} \quad \text{und} \quad \vec{w} = \begin{pmatrix} w_1 \\ w_2 \\ w_3 \end{pmatrix}$$

aus dem Vektorraum k^3 über einem Körper k bezeichnen wir die Zahl

$$\det(\vec{u}, \vec{v}, \vec{w}) \underset{\text{def}}{=} \begin{vmatrix} u_1 & v_1 & w_1 \\ u_2 & v_2 & w_2 \\ u_3 & v_3 & w_3 \end{vmatrix}$$

$$\underset{\text{def}}{=} u_2 v_3 w_1 - u_3 v_2 w_1 + u_3 v_1 w_2 - u_1 v_3 w_2 + u_1 v_2 w_e - u_2 v_1 w_3$$

als *Determinante* von $\vec{u}, \vec{v}$ und $\vec{w}$.

Zumindest im $\mathbb{R}^3$ wissen wir dann, daß drei Vektoren genau dann linear abhängig sind, wenn ihre Determinante verschwindet.

Da der Betrag der Determinante das Volumen des von $\vec{u}, \vec{v}$ und $\vec{w}$ aufgespannten Parallelepipeds ist, ist dieser Betrag unabhängig von der Reihenfolge der Vektoren. Das Vorzeichen hängt aber natürlich von der Reihenfolge ab, denn wegen der Schiefsymmetrie des Vektorprodukts ist beispielsweise

$$\det(\vec{v}, \vec{u}, \vec{w}) = -\det(\vec{u}, \vec{v}, \vec{w})\,;$$

entsprechend ändert sich das Vorzeichen, wie man sich leicht überzeugt, bei jeder Vertauschung der Reihenfolge zweier Vektoren.

Da sowohl das Vektorprodukt als auch das Skalarprodukt linear in ihren Argumenten sind, ist auch det linear in jedem ihrer drei Argumente, wir haben also die drei Eigenschaften

(D1) $\det(\vec{u}, \vec{v}, \vec{w})$ ändert ihr Vorzeichen, nicht aber ihren Betrag, wenn irgendwelche zwei der Argumente miteinander vertauscht werden.

(D2) $\det(\vec{u}, \vec{v}, \vec{w})$ ist linear in jedem ihrer drei Argumente.

(D3) $\det(\vec{u}, \vec{v}, \vec{w}) = 0$ genau dann, wenn $\vec{u}, \vec{v}$ und $\vec{w}$ linear abhängig sind.

c) Forderungen an eine allgemeine Determinante

Wir möchten gerne für beliebige endlichdimensionale Vektorräume V eine Determinantenfunktion definieren, d.h. wir suchen für einen n-dimensionalen Vektorraum V eine Abbildung

$$\det\colon \underbrace{V \times \cdots \times V}_{n \text{ mal}} \to k; \quad (\vec{v}_1, \ldots, \vec{v}_n) \mapsto \det(\vec{v}_1, \ldots, \vec{v}_n)$$

mit den Eigenschaften

(D1) $\det(\vec{v}_1, \ldots, \vec{v}_n)$ ändert ihr Vorzeichen, nicht aber ihren Betrag, wenn irgendwelche zwei ihrer Argumente miteinander vertauscht werden.

(D2) $\det(\vec{v}_1, \ldots, \vec{v}_n)$ ist linear in jedem ihrer n Argumente.

(D3) $\det(\vec{v}_1, \ldots, \vec{v}_n) = 0$ genau dann, wenn $\vec{v}_1, \ldots, \vec{v}_n$ linear abhängig sind.

Um zu zeigen, daß es eine solche Abbildung auch tatsächlich gibt, folgen wir einer auch sonst bei Existenzbeiweisen oftmals nützlichen Strategie: Wir nehmen an, wir *hätten* eine Funktion det mit den Eigenschaften (D1) bis (D3), und versuchen dann, die Funktion aufgrund dieser Eigenschaften möglichst explizit auszurechnen. Dies wird auf eine Formel für det führen, die wir dann als *Definition* nehmen können, wenn wir nachgewiesen haben, daß sie (D1) bis (D3) erfüllt.

Wir nehmen also an, wir hätten eine Funktion det mit den Eigenschaften (D1) bis (D3). Zu ihrer genaueren Untersuchung wählen wir zunächst eine Basis $\{\vec{e}_1, \ldots, \vec{e}_n\}$ von V; nach Eigenschaft (D3) ist dann $\det(\vec{e}_1, \ldots, \vec{e}_n) \neq 0$. Über den genauen Wert können uns die drei Bedingungen nichts sagen, denn mit det erfüllt offensichtlich auch jedes skalare Vielfache $\lambda \cdot \det$ mit $\lambda \neq 0$ die Forderungen (D1) bis (D3).

Wir betrachten nun n beliebige Vektoren $\vec{v}_1, \ldots, \vec{v}_n$ und schreiben diese bezüglich der Basis $\{\vec{e}_1, \ldots, \vec{e}_n\}$ als Linearkombinationen

$$\vec{v}_i = \sum_{j=1}^{n} a_{ij} \vec{e}_j$$

der gewählten Basisvektoren.

Sukzessive Anwendung der Linearitätsregel (D2) auf die n Argumente von det führt auf

$$\begin{aligned}
\det(\vec v_1,\ldots,\vec v_n) &= \det\Big(\sum_{j_1=1}^n a_{1j_1}\vec e_{j_1},\vec v_2,\ldots,\vec v_n\Big)\\
&\underset{(D2)}{=} \sum_{j_1=1}^n a_{1j_1}\det(\vec e_{j_1},\vec v_2,\ldots,\vec v_n)\,.\\
&= \sum_{j_1=1}^n a_{1j_1}\det\Big(\vec e_{j_1},\sum_{j_2=1}^n a_{2j_2}\vec e_{j_2},\vec v_3,\ldots,\vec v_n\Big)\\
&\underset{(D2)}{=} \sum_{j_1=1}^n\sum_{j_2=1}^n a_{1j_1}a_{2j_2}\det(\vec e_{j_1},\vec e_{j_2},\vec v_3,\ldots,\vec v_n)\\
&\vdots\\
&= \sum_{j_1=1}^n\sum_{j_2=1}^n\cdots\sum_{j_n=1}^n a_{1j_1}a_{2j_2}\cdot\cdots\cdot a_{nj_n}\det(\vec e_{j_1},\cdots,\vec e_{j_n})\,.
\end{aligned}$$

Um dies noch weiter ausrechnen zu können, müssen wir die Vektoren $\vec e_{j_\nu}$ in jedem der Summanden in die natürliche Reihenfolge bringen; dabei hilft die Regel (D1). Eine ganze Reihe von Summanden können wir allerdings gleich von vornherein außer Acht lassen, denn nach (D3) ist

$$\det(\vec e_{j_1},\cdots,\vec e_{j_n}) = 0\,,$$

wenn (mindestens) zwei der $\vec e_{j_\nu}$ übereinstimmen: Falls $\vec e_{j_\nu}$ und $\vec e_{j_\mu}$ übereinstimmen, ist $\vec e_{j_\nu} - e_{j_\mu} = \vec 0$ eine nichttriviale Darstellung des Nullvektors. Ein von Null verschiedener Wert ist also nur möglich, wenn $\{j_1,\ldots,j_n\} = \{1,\ldots,n\}$ ist, wenn also die Abbildung

$$\pi\colon\{1,\ldots,n\}\to\{1,\ldots,n\};\quad \nu\mapsto j_\nu$$

bijektiv ist.

Unsere nächste Aufgabe wird daher sein, solche Abbildungen zu studieren und sie auf Vertauschungen zweier Elemente zurückzuführen, so daß wir schließlich die Summanden mittels (D1) auf $\det(\vec e_1,\ldots,\vec e_n)$ zurückführen können.

d) Permutationen

Definition: *a)* Eine Permutation ist eine bijektive Abbildung

$$\pi\colon\{1,\ldots,n\}\to\{1,\ldots,n\}\,.$$

Sie wird auch kurz als

$$\pi = \begin{pmatrix} 1 & 2 & \ldots & n\\ \pi(1) & \pi(2) & \ldots & \pi(n)\end{pmatrix}$$

geschrieben.

b) Eine Transposition ist eine Permutation τ, die zwei Elemente von $\{1,\ldots,n\}$ vertauscht und den Rest festläßt. Sind i,j die beiden Elemente, die von τ vertauscht werden, so schreiben wir kurz $\tau = (i\ j)$.

c) Die Menge aller Permutationen $\pi\colon\{1,\ldots,n\}\to\{1,\ldots,n\}$ wird als *symmetrische Gruppe* $\mathfrak{S}_n$ bezeichnet.

Die Matrixdarstellung einer Permutation hat natürlich nichts mit Abbildungsmatrizen von linearen Abbildungen zu tun; hier verwenden wir die Matrix nur als Darstellung der Wertetabelle von π.

Das Wort *Gruppe* tauchte bereits in §1 auf im Zusammenhang mit der Definition eines Vektorraum; es wird langsam Zeit, es wirklich zu definieren. Intuitiv versteht man unter einer Gruppe eine Struktur, deren Elemente sich so miteinander verknüpfen lassen, daß die „üblichen" Rechenregeln gelten – mit Ausnahme, eventuell, des Kommutativgesetzes, denn dieses gilt ja beispielsweise schon bei der Matrixmultiplikation nicht mehr. Die exaktes Definition ist die folgende:

Definition: Eine *Gruppe* $(G,\circ)$ ist eine Menge G zusammen mit einer inneren Verknüpfung $\circ\colon G\times G\to G$, für die gilt:

1. $g\circ(h\circ k) = (g\circ h)\circ k$ für alle $g,h,k\in G$
2. Es gibt ein *Neutralelement* $e\in G$, so daß für alle $g\in G$ gilt $e\circ g = g\circ e = g$.
3. Zu jedem Element $g\in G$ gibt es ein *inverses Element* $g'\in g$, so daß $g\circ g' = g'\circ g = e$ ist.

Die Gruppe heißt *kommutativ* oder *abelsch,* wenn zusätzlich gilt

4. $g \circ h = h \circ g$ für alle $g, h \in G$.

In diesem Sinne bilden beispielsweise die ganzen Zahlen mit der Addition als Verknüpfung eine Gruppe, sogar eine abelsche Gruppe, nicht aber die natürlichen Zahlen, denn es gibt beispielsweise keine natürliche Zahl n, so daß $n + 2 = 0$ ist, und auch die Null liegt zumindest nach der in dieser Vorlesung zugrundegelegten Konvention nicht in $\mathbb{N}$. Entsprechend bilden die ganzen Zahlen bezüglich der *Multiplikation* keine Gruppe, da die Gleichung $5x = 1$ in $\mathbb{Z}$ nicht lösbar ist, aber die positiven rationalen Zahlen bilden eine multiplikative Gruppe. Die rationalen Zahlen bilden keine, da $0x = 1$ unlösbar ist, aber die rationalen Zahlen ohne Null sind eine multiplikative Gruppe, genauso auch die *invertierbaren* $n \times n$-Matrizen.

Kommutative Gruppen werden meist als *abelsch* bezeichnet zu Ehren des norwegischen Mathematikers NILS HENRIK ABEL (1802–1829), der bei seinen Untersuchungen über Integrale algebraischer Funktionen auf solche Gruppen stieß. Diese heute nach ihm benannten *abelschen Integrale* treten beispielsweise auf bei der Berechnung der Bogenlänge auf einer Ellipse; die von ABEL gefundenen Gruppen spielen heute eine wichtige Rolle in der Kryptographie. 1824 bewies er mit Hilfe von Permutationen (die, wie wir gleich sehen werden, *nichtabelsche* Gruppen bilden), daß Gleichungen fünften Grades im allgemeinen nicht durch Wurzelausdrücke gelöst werden können.

Da wir die Menge aller Permutationen als *symmetrische Gruppe* bezeichnen, sollten auch diese eine Gruppe bilden.

Die Gruppenoperation ist natürlich die Hintereinanderausführung von Abbildungen, das Produkt zweier Permutationen π und ω ist also die Abbildung $\pi \circ \omega$, die eine Zahl i abbildet auf $(\pi \circ \omega)(i) = \pi\big(\omega(i)\big)$.

Die Assoziativität der Verknüpfung ist, wie stets bei der Hintereinanderausführung von Abbildungen, klar; Einselement ist die identische Permutation, und Inverse gibt es, da eine Permutation nach Definition bijektiv ist und somit eine Umkehrabbildung hat. Speziell für Transpositionen ist die Bestimmung dieser Umkehrabbildung besonders einfach:

Da eine Transposition nicht anderes tut, als zwei Elemente miteinander zu vertauschen, macht sie sich selbst rückgängig und ist somit ihr eigenes Inverses.

Wie meist bei der Hintereinanderausführung von Abbildungen ist auch bei Permutationen das Kommutativgesetz nicht erfüllt. Als Beispiel betrachten wir die beiden Transpositionen $(1\ 2)$ und $(1; 3)$.

Beim Produkt $(1\ 2) \circ (1\ 3)$ wird zuerst die Transposition $(1\ 3)$ ausgeführt, die die Zahlen 1 und 3 vertauscht, sodann vertauscht $(1\ 2)$ die Zahlen 1 und 2:

$$\begin{array}{ccc} 1 & 2 & 3 \\ \downarrow & \downarrow & \downarrow \\ 3 & 2 & 1 \\ \downarrow & \downarrow & \downarrow \\ 3 & 1 & 2 \end{array}$$

Somit ist

$$(1\ 2) \circ (1\ 3) = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix} .$$

Beim Produkt $(1\ 3) \circ (1\ 2)$ dagegen wird zuerst die Transposition $(1\ 3)$ ausgeführt und dann erst $(1\ 2)$; hier ist der Gang der Vertauschungen also

$$\begin{array}{ccc} 1 & 2 & 3 \\ \downarrow & \downarrow & \downarrow \\ 2 & 1 & 3 \\ \downarrow & \downarrow & \downarrow \\ 2 & 3 & 1\,, \end{array}$$

d.h.

$$(1\ 3) \circ (1\ 2) = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix} .$$

Aus der Matrixdarstellung

$$\begin{pmatrix} 1 & 2 & \dots & n \\ \pi(1) & \pi(2) & \dots & \pi(n) \end{pmatrix}$$

einer Permutation läßt sich leicht die Anzahl möglicher Permutationen von n Elementen ablesen: Füllen wir die untere Zeile von links aus systematisch auf, so gibt es für $\pi(1)$ noch die volle Auswahl unter allen

n Elementen, für $\pi(2)$ kommen alle Elemente in Frage *außer* $\pi(1)$, also nur noch $n-1$ Stück, für $\pi(3)$ gibt es entsprechend nur noch $n-2$, bis es schließlich bei $\pi(n)$ überhaupt nichts mehr zu wählen gibt, denn $\pi(n)$ ist die einzige Zahl, die bis dahin noch nicht in der zweiten Zeile der Matrix steht. Insgesamt gibt es also

$$n \cdot (n-1) \cdot \cdots \cdot 2 \cdot 1 = n!$$

Permutationen von $\{1, \ldots, n\}$, die symmetrische Gruppe $\mathfrak{S}_n$ enthält somit $n!$ Elemente.

Für die Anwendung von Regel (D1) auf die weitere Berechnung von Determinanten ist der folgende Satz wichtig:

Satz: Jede Permutation π kann als Produkt

$$\pi = \tau_1 \circ \cdots \circ \tau_r$$

von Transpositionen geschrieben werden.

Beweis: Da es sich hier um eine Mathematikvorlesung für (technische) Informatiker handelt, möchte ich sowohl einen mathematischen als auch einen informatischen Beweis geben.

Der einfachste mathematische Beweis benutzt vollständige Induktion nach der Elementanzahl n der zu permutierenden Menge. Für $n = 1$ gibt es keine Permutation außer der Identität, die man als leeres Produkt von Transpositionen betrachten kann; da aber nicht jeder diese Art von logischen Taschenspielertricks liebt, betrachten wir zur Vorsicht auch noch den Fall $n = 2$ als weitere Verankerung der Induktion.

Hier gibt es genau zwei Permutationen: Die Identität, die alles festläßt, und die Vertauschung der beiden Elemente. Letzere ist eine Transposition, erstere kann wieder als leeres Produkt von Transpositionen betrachtet werden oder, falls man das nicht möchte, als Quadrat der Vertauschung, denn nach zweimaligem Vertauschen ist wieder alles beim alten.

Beim *Induktionsschritt* nehmen wir an, wir hätten den Satz für Permutationen $(n-1)$-elementiger Mengen bewiesen und betrachten nun eine Permutation

$$\pi\colon \{1, \ldots, n\} \to \{1, \ldots, n\}.$$

Für diese gibt es ein Element $i \in \{1, \ldots, n\}$, so daß $\pi(i) = n$ ist, denn π ist ja bijektiv. Die Permutation

$$\pi' = \pi \circ (i\ n)$$

bildet n ab auf

$$\pi'(n) = \big(\pi \circ (i\ n)\big)(n) = \pi\big((i\ n)(n)\big) = \pi(i) = n\,,$$

d.h. n bleibt fest. Da π' bijektiv ist, muß es daher auch die Menge $\{1, \ldots, n-1\}$ auf sich selbst abbilden; wenn wir n ignorieren, können wir π' daher auch als eine Permutation von $\{1, \ldots, n-1\}$ auffassen. Von der wissen wir, daß sie als Produkt

$$\pi' = \tau_1 \circ \cdots \circ \tau_r$$

von Transpositionen darstellbar ist. Wegen $\pi' = \pi \circ (i\ \ n)$ und weil Transpositionen zu sich selbst invers sind, ist damit auch

$$\pi = \pi' \circ (i\ n) = \tau_1 \circ \cdots \circ \tau_r \circ (i\ n)$$

als Produkt von Transpositionen darstellbar, und damit ist der Satz einmal bewiesen.

Zum zweiten Beweis, auf Grundlage der Informatik, führt die Beobachtung, daß jeder gängige Sortieralgorithmus, der auf Vertauschungen von Elementen beruht (und das tut fast jeder) ein konstruktives Verfahren liefert, um Permutationen als Produkte von Transpositionen zu schreiben: Soll π zerlegt werden, so sortiere man durch fortgesetztes Vertauschungen zweier Zahlen die Folge

$$\pi(1),\ \pi(2),\ \pi(3),\ \ldots, \pi(n-1),\ \pi(n)$$

der Größe nach; die erste angewandte Transposition sei τ_1, die letzte τ_n.

Die sortierte Folge der $\pi(i)$ ist natürlich $1, \ldots, n$ ist; daher für jedes i

$$(\tau_r \circ \cdots \circ \tau_1 \circ \pi)(i) = i\,,$$

d.h.

$$\tau_r \circ \cdots \circ \tau_1 \circ \pi = \text{Identität}\,.$$

Multipliziert man beide Seiten von links mit $\tau_1 \circ \cdots \circ \tau_r$, so folgt

$$\pi = \tau_1 \circ \cdots \circ \tau_r\,,$$

denn

$$\begin{aligned}
&(\tau_1 \circ \cdots \circ \tau_r) \circ (\tau_r \circ \cdots \circ \tau_1) \\
&= \tau_1 \circ \cdots \circ \tau_{r-1} \circ (\tau_r \circ \tau_r) \circ \tau_{r-1} \circ \cdots \circ \tau_1 \\
&= \tau_1 \circ \cdots \circ \tau_{r-2} \circ (\tau_{r-1} \circ \tau_{r-1}) \circ \tau_{r-2} \circ \cdots \circ \tau_1 \\
&= \tau_1 \circ \cdots \circ \tau_{r-3} \circ (\tau_{r-2} \circ \tau_{r-2}) \circ \tau_{r-3} \cdots \circ \tau_1 \\
&= \cdots = \tau_1 \circ (\tau_2 \circ \tau_2) \circ \tau_1 = \tau_1 \circ \tau_1 = \text{Identität}\,,
\end{aligned}$$

Also ist π als Produkt der Transpositionen τ_1 bis τ_r darstellbar. ■

Ein Leser, der mit den gebräuchlichen Sortierverfahren vertraut ist, wird unschwer erkennen, daß der „mathematische“ Beweis ein Spezialfall des informatischen ist, in dem als Sortierverfahren ein relativ einfaches $O(n^2)$-Verfahren benutzt wurde (das allerdings zumindest für einstellige n den meisten „besseren“ Sortierverfahren überlegen sein dürfte). Die beiden Beweise unterscheiden sich also nicht sonderlich.

Die Anzahl n der benötigten Vertauschungen hängt natürlich von der Vorgehensweise und der Effiziens des gewählten Sortierverfahrens ab; die Darstellung von π als Produkt von Transpositionen ist also alles andere als eindeutig. Da wir für die Determinantenberechnung vor allem wissen müssen, ob die Abzahl der Vertauschungen gerade oder ungerade ist, wollen wir uns als nächstes davon überzeugen, daß wenigstens dies, d.h. also die Parität der Anzahl r der benötigten Transpositionen, eindeutig bestimmt ist: r ist für gegebenes π entweder immer gerade oder immer ungerade.

Dazu definieren wir für jede Permutation $\pi \in \mathfrak{S}_n$ ein Vorzeichen:

Definition: Das *Vorzeichen* der Permutation $\pi \in \mathfrak{S}_n$ ist die Zahl

$$\varepsilon(\pi) \underset{\text{def}}{=} \prod_{\{i,j\}\subset\{1,\ldots,n\}} \frac{\pi(j)-\pi(i)}{j-i}\,,$$

wobei das Produkt über die sämtlichen zweielementigen Teilmengen von $\{1,\ldots,n\}$ genommen wird.

$\varepsilon(\pi)$ ist wirklich ein Vorzeichen, das heißt gleich ± 1, denn

$$|\varepsilon(\pi)| = \frac{\prod\limits_{\{i,j\}\subset\{1,\ldots,n\}} |\pi(j)-\pi(i)|}{\prod\limits_{\{i,j\}\subset\{1,\ldots,n\}} |j-i|} = 1\,,$$

weil mit $\{i,j\}$ wegen der Bijektivität von π auch $\{\pi(i),\pi(j)\}$ die sämtlichen zweielementigen Teilmengen von $\{1,\ldots,n\}$ durchläuft, so daß im Zähler und Nenner bis auf Reihenfolge genau dieselben Faktoren stehen.

Für eine Transposition $\tau = (k\ \ell)$ ist

$$\frac{\tau(j)-\tau(i)}{j-i} = \frac{j-i}{j-i} = 1 \quad \text{falls} \quad \{i,j\}\cap\{k,l\} = \emptyset\,,$$

da $(k\ \ell)$ alle Zahlen außer k und ℓ festläßt. Die Zahlen k und ℓ werden vertauscht, daher ist

$$\frac{\tau(k)-\tau(\ell)}{k-\ell} = \frac{\ell-k}{k-\ell} = -1\,.$$

Bleibt noch der Fall, daß eine der beiden Zahlen i oder j gleich k oder ℓ ist; da wir zweielementige *Mengen* betrachten und $\{i,j\} = \{j,i\}$ ist, können wir vereinbaren, daß diese Zahl i sein soll.

Ist $i = k$, so ist das Produkt der Terme für $\{k,j\}$ und $\{\ell,j\}$ gleich

$$\frac{\tau(k)-\tau(j)}{k-j}\cdot\frac{\tau(\ell)-\tau(j)}{\ell-j} = \frac{\ell-j}{k-j}\cdot\frac{k-j}{\ell-j} = +1\,;$$

und genau entsprechend kann man auch für $i = \ell$ argumentieren.

Bildet man daher das Produkt über alle zweielementigen Teilmengen von $\{1,\ldots,n\}$, so erhält man $\varepsilon(\tau) = -1$ und damit das

Lemma: Das Vorzeichen einer Transposition ist -1. ■

Der Zusammenhang zwischen dem Vorzeichen und der Transpositionsdarstellung einer Permutation ergibt sich aus

Lemma: Für zwei Permutationen $\pi,\pi' \in \mathfrak{S}_n$ ist

$$\varepsilon(\pi\circ\pi') = \varepsilon(\pi)\cdot\varepsilon(\pi')\,.$$

Beweis: Nach Definition ist

$$\begin{aligned}\varepsilon(\pi\circ\pi') &= \prod_{\{i,j\}\subset\{1,\ldots,n\}} \frac{(\pi\circ\pi')(j)-(\pi\circ\pi')(i)}{j-i}\\ &= \prod_{\{i,j\}\subset\{1,\ldots,n\}} \frac{\pi\bigl(\pi'(j)\bigr)-\pi\bigl(\pi'(i)\bigr)}{j-i}\,.\end{aligned}$$

Nach Erweiterung mit

$$\prod_{\{i,j\}\subset\{1,\ldots,n\}} \frac{\pi'(j)-\pi'(i)}{\pi'(j)-\pi'(i)}$$

und Umordnung wird dies zu

$$\varepsilon(\pi\circ\pi') = \prod_{\{i,j\}} \frac{\pi\bigl(\pi'(j)\bigr)-\pi\bigl(\pi'(i)\bigr)}{\pi'(j)-\pi'(i)} \cdot \prod_{\{i,j\}} \frac{\pi'(j)-\pi'(i)}{j-i}\,.$$

Das zweite dieser Produkte ist natürlich $\varepsilon(\pi')$, und das erste ist $\varepsilon(\pi)$, da mit $\{i,j\}$ auch $\{\pi'(i),\pi'(j)\}$ die sämtlichen zweielementigen Teilmengen von $\{1,\ldots,n\}$ durchläuft. ∎

Sind nun

$$\pi = \tau_1\circ\cdots\circ\tau_r = \sigma_1\circ\cdots\circ\sigma_s$$

zwei Darstellungen einer Permutation π als Produkt von Transpositionen, so ist nach den gerade bewiesenen Lemmata einerseits

$$\varepsilon(\pi) = \prod_{i=1}^{r}\varepsilon(\tau_i) = (-1)^r,$$

andererseits aber auch

$$\varepsilon(\pi) = \prod_{i=1}^{s}\varepsilon(\sigma_i) = (-1)^s,$$

also muß $(-1)^r = (-1)^s$ sein und damit r, s entweder beide gerade oder beide ungerade.

Definition: Eine Permutation $\pi\in\mathfrak{S}_n$ heißt *gerade,* wenn sie als Produkt einer geraden Anzahl von Transpositionen geschrieben werden kann, ansonsten heißt sie *ungerade.*

Das Vorzeichen $\varepsilon(\pi)$ von π ist also +1 für gerades π und -1 für ungerades.

e) Existenz von Determinanten

Nach diesem Einschub über Permutationen und Transpositionen haben wir das nötige Rüstzeug, um die Rechnung aus Abschnitt *b)* zu Ende zu führen:

Da die Determinante bei jeder Vertauschung zweier Elemente ihr Vorzeichen wechselt, folgt sofort aus den obigen Betrachtungen, daß für jede Permutation $\pi\in\mathfrak{S}_n$ gilt

$$\det\bigl(\vec e_{\pi(1)},\ldots,\vec e_{\pi(n)}\bigr) = \begin{cases} +\det(\vec e_1,\ldots,\vec e_n) & \text{falls } \pi \text{ gerade}\\ -\det(\vec e_1,\ldots,\vec e_n) & \text{falls } \pi \text{ ungerade}\end{cases},$$

d.h. $\det\bigl(\vec e_{\pi(1)},\ldots,\vec e_{\pi(n)}\bigr) = \varepsilon(\pi)\cdot\det(\vec e_1,\ldots,\vec e_n)$. Wenn wir die Rechnung aus Abschnitt *b)* fortführen, erhalten wir daher

$$\begin{aligned}\det\bigl(\vec v_1,\ldots,\vec v_n\bigr) &= \sum_{\pi\in\mathfrak{S}_n} a_{1\pi(1)}\cdots a_{n\pi(n)}\cdot\det\bigl(\vec e_{\pi(1)},\ldots,\vec e_{\pi(n)}\bigr)\\ &= \left(\sum_{\pi\in\mathfrak{S}_n}\varepsilon(\pi)a_{1\pi(1)}\cdots a_{n\pi(n)}\right)\cdot\det\bigl(\vec e_1,\ldots,\vec e_n\bigr)\,.\end{aligned}$$

Somit ist $\det\bigl(\vec v_1,\ldots,\vec v_n\bigr)$ eindeutig bestimmt bis auf den Wert von $\det\bigl(\vec e_1,\ldots,\vec e_n\bigr)$; falls es überhaupt Determinanten gibt, müssen sie also so aussehen. Wir definieren daher

Definition: Die Determinante von n Vektoren

$$\vec v_1 = a_{11}\vec e_1+\cdots+a_{1n}\vec e_n,\quad \ldots,\quad \vec v_n = a_{n1}\vec e_1+\cdots+a_{nn}\vec e_n$$

des n-dimensionalen Vektorraums V bezüglich der Basis $\vec e_1,\ldots,\vec e_n$ ist

$$\det\bigl(\vec v_1,\ldots,\vec v_n\bigr) = \sum_{\pi\in\mathfrak{S}_n}\varepsilon(\pi)a_{1\pi(1)}\cdots a_{n\pi(n)}\,.$$

Insbesondere ist also $\det(\vec e_1,\ldots,\vec e_n) = 1$.

Satz: Die so definierte Determinante hat die Eigenschaften (D1) bis (D3).

Beweis: Beginnen wir mit (D1), d.h. die Determinante wechselt ihr Vorzeichen, wenn zwei der Argumente vertauscht werden. Vertauscht man etwa die Argumente $\vec{v}_i$ und $\vec{v}_j$, so werden anstelle der Summanden

$$\varepsilon(\pi)\, a_{1\pi(1)} \cdots a_{i\pi(i)} \cdots a_{j\pi(j)} \cdots a_{n\pi(n)}$$

die Summanden

$$\varepsilon(\pi)\, b_{1\pi(1)} \cdots b_{j\pi(j)} \cdots b_{i\pi(i)} \cdots b_{n\pi(n)}$$

aufaddiert, wobei

$$b_{k\ell} = \begin{cases} a_{k\ell} & \text{falls } k \neq i, j \\ a_{j\ell} & \text{falls } k = i \\ a_{i\ell} & \text{falls } k = j \end{cases}$$

ist. Die Summanden sind also

$$\begin{aligned} &\varepsilon(\pi)\, a_{1\pi(1)} \cdots a_{j\pi(i)} \cdots a_{i\pi(j)} \cdots a_{n\pi(n)} \\ =\ &\varepsilon(\pi)\, a_{1\pi(1)} \cdots a_{i\pi(j)} \cdots a_{j\pi(i)} \cdots a_{n\pi(n)} \\ =\ &\varepsilon(\pi)\, a_{1\pi'(1)} \cdots a_{i\pi'(i)} \cdots a_{j\pi'(j)} \cdots a_{n\pi'(n)} \end{aligned}$$

mit

$$\pi'(k) = \begin{cases} \pi(k) & \text{falls } k \neq i, j \\ \pi(j) & \text{falls } k = i \\ \pi(i) & \text{falls } k = j\ , \end{cases}$$

d.h. $\pi' = \pi \circ (i\ j)$. Als Produkt von Transpositionen geschrieben hat π' also einen Faktor mehr als π; falls π ungerade war, ist daher π' gerade, und umgekehrt. Somit ist

$$\varepsilon(\pi') = -\varepsilon(\pi)$$

und damit lassen sich die Summanden auch schreiben als

$$-\varepsilon(\pi')\, a_{1\pi'(1)} \cdots a_{i\pi'(i)} \cdots a_{j\pi'(j)} \cdots a_{n\pi'(n)}\ .$$

Summiert man über alle $\pi' \in \mathfrak{S}_n$, ist also

$$\begin{aligned} &\det(\vec{v}_1, \ldots, \vec{v}_j, \ldots, \vec{v}_i, \ldots, \vec{v}_n) \\ = &\sum_{\pi' \in \mathfrak{S}_n} \bigl(-\varepsilon(\pi')\bigr)\, a_{1\pi'(1)} \cdots a_{i\pi'(i)} \cdots a_{j\pi'(j)} \cdots a_{n\pi'(n)} \\ = &- \sum_{\pi' \in \mathfrak{S}_n} \varepsilon(\pi')\, a_{1\pi'(1)} \cdots a_{i\pi'(i)} \cdots a_{j\pi'(j)} \cdots a_{n\pi'(n)} \\ = &- \det(\vec{v}_1, \ldots, \vec{v}_i, \ldots, \vec{v}_j, \ldots, \vec{v}_n)\ , \end{aligned}$$

womit (D1) bewiesen wäre.

Eigenschaft (D2), die Linearität in jedem der Argumente, ist klar, denn jedes der Monome $a_{1\pi(1)} \cdots a_{n\pi(n)}$ enthält *genau eine* Komponente des Vektors $\vec{v}_i$ und ist damit linear in $\vec{v}_i$.

Bleibt noch (D3), d.h. $\det\bigl(\vec{v}_1, \ldots, \vec{v}_n\bigr)$ soll genau dann verschwinden, wenn die $\vec{v}_i$ linear abhängig sind.

Sind zunächst die Vektoren $\vec{v}_1, \ldots, \vec{v}_n$ linear abhängig, so läßt sich einer von ihnen als Linearkombination der anderen schreiben; durch Umordnen (wobei die Determinante nach der schon bewiesenen Eigenschaft (D1) höchstens ihr Vorzeichen ändert) können wir annehmen, daß dies etwa $\vec{v}_1$ sei; wir setzen

$$\vec{v}_1 = \sum_{i=2}^{n} \lambda_i \vec{v}_i\ .$$

Dann ist

$$\begin{aligned} \det\bigl(\vec{v}_1, \ldots, \vec{v}_n\bigr) &= \det\left(\sum_{i=2}^{n} \lambda_i \vec{v}_i, \vec{v}_2, \ldots, \vec{v}_n\right) \\ &\underset{(D2)}{=} \sum_{i=2}^{n} \lambda_i \det\bigl(\vec{v}_i, \vec{v}_2, \ldots, \vec{v}_n\bigr)\ . \end{aligned}$$

Im i-ten Summanden der letzten Summe tritt der Vektor $\vec{v}_i$ zweimal als Argument von det auf: Einmal an der ersten, und dann noch an der i-ten. Vertauscht man diese beiden (gleichen) Argumente, ändert sich natürlich nichts; andererseits haben wir aber bereits (D1) bewiesen, wonach sich beim Vertauschen irgend zweier Argumente das Vorzeichen ändert. Der i-te Summand hat also einen Wert s_i, für den $s_i = -s_i$ ist. Falls wir mit reellen Zahlen arbeiten, folgt hieraus natürlich sofort, daß alle s_i verschwinden und somit auch ihre Summe.

Für beliebige Körper gilt das aber leider nicht: Aus $s_i = -s_i$ folgt, wenn wir auf beiden Seiten s_i addieren, zunächst nur, daß

$$s_i + s_i = (1+1)s_i = 0$$

ist, und daraus folgt genau dann, daß auch $s_i = 0$ ist, wenn $1+1 \neq 0$ ist. Dies ist aber beispielweise in dem für die digitale Informationsverarbeitung wichtigen Körper $\mathbb{F}_2 = \{0,1\}$ nicht der Fall; wir müssen uns also mit Rücksicht auf diesen (und eine ganze Reihe anderer) Körper noch überlegen, daß wirklich immer

$$\det(\vec{v}_1, \ldots, \vec{v}_n) = 0$$

ist, wenn zwei der Argumente $\vec{v}_i$ übereinstimmen.

Dazu beachten wir, daß im Falle $\vec{v}_i = \vec{v}_j$ auch für alle ℓ die Koeffizienten $a_{i\ell}$ und $a_{j\ell}$ übereinstimmen, d.h. in der Summe

$$\det A = \sum_{\pi \in \mathfrak{S}_n} \varepsilon(\pi) a_{1\pi(1)} \cdots a_{i\pi(i)} \cdots a_{j\pi(j)} \cdots a_{n\pi(n)}$$

ist stets

$$a_{1\pi(1)} \cdots a_{i\pi(i)} \cdots a_{j\pi(j)} \cdots a_{n\pi(n)} = a_{1\pi(1)} \cdots a_{i\pi(j)} \cdots a_{j\pi(i)} \cdots a_{n\pi(n)} .$$

Hier ist die rechte Seite gerade der Summand zur Permutation $\pi \circ (i\ j)$, die genau dann gerade ist, wenn π ungerade ist, und umgekehrt. Also haben beide Terme verschiedenes Vorzeichen und heben sich gegenseitig weg; die Summe über alle Permutationen ist also Null.

Also verschwindet die Determinante unabhängig vom Grundkörper immer dann, wenn die Vektoren linear abhängig sind.

Nun seien $\vec{v}_1, \ldots, \vec{v}_n$ linear unabhängig; für dieses Fall müssen wir zeigen, daß die Determinante *nicht* verschwindet. Wir wissen, daß für die Basisvektoren $\vec{e}_1, \ldots, \vec{e}_n$

$$\det(\vec{e}_1, \ldots, \vec{e}_n) = 1 \neq 0$$

ist; außerdem wissen wir, daß n beliebige linear unabhängige Vektoren eines n-dimensionalen Vektorraums eine Basis bilden.

Daher können wir die $\vec{e}_i$ als Linearkombinationen der $\vec{v}_j$ schreiben, etwa als

$$\vec{e}_i = \sum_{j=1}^{n} b_{ij} \vec{v}_j .$$

Nun können wir die Rechnungen des letzten Paragraphen wörtlich wiederholen, nur daß dieses Mal $\vec{v}_1, \ldots, \vec{v}_n$ die Basisvektoren sind und $\vec{e}_1, \ldots, \vec{e}_n$ die Vektoren, für die wir $\det$ berechnen möchten. Wir wissen zwar noch nicht, daß $\det(\vec{e}_1, \ldots, \vec{e}_n)$ nicht verschwindet, aber das haben wir in der dortigen Rechnung auch nie verwendet; alles hing nur ab von den bereits bewiesenen Forderungen (D1) und (D2) und der ebenfalls schon bewiesenen Tatsache, daß $\det$ verschwindet, wenn zwei der Argumente gleich sind. Somit erhalten wir die Formel

$$\det(\vec{e}_1, \ldots, \vec{e}_n) = \left(\sum_{\pi \in \mathfrak{S}_n} \varepsilon(\pi) b_{1\pi(1)} \cdots b_{n\pi(n)} \right) \cdot \det(\vec{v}_1, \ldots, \vec{v}_n) .$$

In dieser Gleichung steht links die Zahl eins, also kann keiner der beiden Faktoren auf der rechten Seite verschwinden; insbesondere ist $\det(\vec{v}_1, \ldots, \vec{v}_n) \neq 0$.

Damit ist der Satz vollständig bewiesen. ■

Fassen wir zusammen, was wir bislang in diesem Paragraphen gemacht haben:

- Wir haben anhand des Beispiels von Determinanten im $\mathbb{R}^2$ und $\mathbb{R}^3$ drei zentrale Forderungen an eine allgemeine Determinante aufgestellt.
- Unter der Annahme, daß es Funktionen gibt, die diese drei Forderungen erfüllen, haben wir ausgerechnet, wie solche Funktionen aussehen müssen; inbsbesondere haben wir gesehen, daß sie durch ihren Wert auf den Vektoren einer Basis eindeutig bestimmt sind.
- Dann haben wir die so erhaltene explizite Formel als *Definition* einer allgemeinen Determinanten genommen und gezeigt, daß die so definierte Funktion tatsächlich die drei Forderungen erfüllt.

Insgesamt haben wir also gesehen, daß es Funktionen gibt, die den drei Forderungen (D1) bis (D3) genügen und daß solche Funktionen bis auf eine multiplikative Konstante eindeutig bestimmt sind.

f) Die Determinante einer Matrix

Definition: Unter der Determinante einer Matrix $A \in k^{n\times n}$ verstehen wir die Determinante ihrer Spaltenvektoren; ist also

$$A = (a_{ij}) \quad \text{und} \quad \vec{v}_j = \begin{pmatrix} a_{1j} \\ \vdots \\ a_{nj} \end{pmatrix},$$

so ist

$$\det A \underset{\text{def}}{=} \det\bigl(\vec{v}_1, \dots, \vec{v}_n\bigr) .$$

Wir schreiben auch kurz

$$\det A = |A| = \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix} .$$

Um etwas Übung im Umgang mit Determinanten zu bekommen, wollen wir dies für $n = 1$ und die beiden bereits bekannten Fälle $n = 2$ und $n = 3$ noch einmal explizit ausrechnen – auch wenn wir bereits wissen, was herauskommen muß.

Der Fall $n = 1$ einer 1×1-Matrix ist völlig uninteressant: Die einzige Permutation einer einelementigen Menge ist die Identität, die für jedes n gerade ist; für $A = (a) \in k^{1\times 1}$ ist also $\det A = a$. Die Schreibweise $\det A = |a|$ ist hier irreführend, da es zumindest für $k = \mathbb{R}$ oder $\mathbb{C}$ zu Verwechslungen mit der Betragsfunktion kommen kann. Da Determinanten von 1×1-Matrizen völlig uninteressant sind, ist dies jedoch nicht weiter schlimm.

Für $n = 2$ gibt es zwei Permutationen: Die (gerade) Identität und die (ungerade) Transposition (1 2). Nach Definition der (allgemeinen) Determinanten ist

$$\det \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}a_{21} ,$$

wie erwartet.

Für $n = 3$ gibt es bereits $3! = 6$ Permutationen, darunter natürlich die (gerade) Identität.

Jede weitere Permutation kann höchstens ein Element festlassen, denn würde sie zwei festlassen, müßte wegen der Bijektivität der Abbildung auch das dritte auf sich selbst abgebildet werden, wir hätten also die Identität.

Falls ein Element festgehalten wird, bleibt einer nichtidentischen Permutation daher nichts anderes übrig, als die beiden anderen zu vertauschen; dies führt auf die drei (ungeraden) Transpositionen (2 3), (1 3) und (1 2).

Falls kein Element auf sich selbst abgebildet wird, muß die Eins entweder auf 2 oder auf 3 abgebildet werden. Im ersten Fall geht dann 2 auf 3, denn sonst müßte zwei entweder festbleiben oder wir hätten die Transposition (1 2), die drei auf sich selbst abbildet. Wegen der Bijektivität der Abbildung gibt es dann keine andere Möglichkeit als daß drei auf eins abgebildet wird; wir haben also die zyklische Vertauschung

$$1 \mapsto 2 \mapsto 3 \mapsto 1 .$$

Diese Permutation ist gerade, denn sie läßt sich beispielsweise als Produkt (1 3)(1 2) zweier Transpositionen schreiben: Wenden wir dieses Produkt an auf 1, so wird die Eins zunächst von (1 2) auf zwei abgebildet, und zwei bleibt fest unter (1 3), so daß insgesamt eins auf zwei abgebildet wird, wie verlangt.

Zwei wird von (1 2) auf die Eins abgebildet, und die wiederum von (1 3) auf drei, also haben wir auch hier insgesamt das richtige Ergebnis. Drei schließlich kann dann wegen der Bijektivität von (1 3)(1 2) nur auf eins abgebildet werden, was man auch schnell direkt sieht, denn (1 2) läßt die Drei unverändert, während (1 3) sie auf eins abbildet.

Die noch verbleibende Permutation bildet eins auf drei und daher drei auf zwei ab, sie ist also die zyklische Vertauschung

$$1 \mapsto 3 \mapsto 2 \mapsto 1 \quad \text{oder} \quad 3 \mapsto 2 \mapsto 1 \mapsto 3 ,$$

d.h. die Umkehrabbildung zur gerade betrachteten Permutation. Daher ist sie als Produkt von Transpositionen gleich

$$\big((1\ 2)(1\ 3)\big)^{-1} = (1\ 3)^{-1}(1\ 2)^{-1} = (1\ 3)(1\ 2)$$

und somit auch gerade.

Zur Berechnung der Determinanten einer 3×3-Matrix (a_{ij}) beginnen wir mit den geraden Permutationen: Die Identität liefert den Term $a_{11}a_{22}a_{33}$, d.h. das Produkt der drei Diagonalelemente, die zyklische Vertauschung $1 \to 2 \to 3 \to 1$ liefert das Produkt $a_{12}a_{23}a_{31}$, und ihr Inverses $a_{13}a_{21}a_{32}$.

Die drei Transpositionen führen auf die negativ zu nehmenden Produkte $a_{11}a_{23}a_{32}$, $a_{13}a_{22}a_{31}$ und $a_{12}a_{21}a_{33}$, insgesamt ist also

$$\begin{aligned}\det A = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} &= a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} \\ &\quad - a_{11}a_{23}a_{32} - a_{13}a_{22}a_{31} - a_{12}a_{21}a_{33}\,,\end{aligned}$$

genau wie wir es in Abschnitt *b)* über das Spatprodukt berechnet hatten und uns aufgrund der SARRUSschen Regel leicht merken können.

Für die Determinanten von 4×4-Matrizen gibt es keine entsprechende Regel mehr. Da nun bereits $4! = 24$ Summanden berücksichtigt werden müssen, eignet sich die Formel aus der Definition der Determinanten nicht mehr gut zur Berechnung, von noch größeren Matrizen ganz zu schweigen: Wie aus der *Analysis I* oder der *Praktischen Informatik* bekannt, ist $\log n! \approx n \log n$, d.h. die Anzahl $n!$ der Summanden wächst mit größer werdendem n noch schneller als die Exponentialfunktion.

Wir werden im folgenden daher auf Verfahren hinarbeiten, die es erlauben, auch größere Determinanten mit vertretbarem Aufwand zu berechnen.

Wir kennen schon einige Rechenregeln für Determinanten als Funktionen von n Vektoren, beispielsweise die Forderungen (D1) bis (D3), aus denen wir die Determinantendefinition hergeleitet haben oder auch die gerade angewandte Regel

$$\begin{aligned}&\det\big(\vec v_1, \dots, \vec v_i + \lambda \vec v_j, \dots, \vec v_n\big) \\ =\, &\det\big(\vec v_1, \dots, \vec v_i, \dots, \vec v_n\big) + \lambda \det\big(\vec v_1, \dots, \vec v_j, \dots, \vec v_n\big) \\ =\, &\det\big(\vec v_1, \dots, \vec v_i, \dots, \vec v_n\big)\,.\end{aligned}$$

In diesem Abschnitt soll es nun um Rechenregeln für Determinanten von Matrizen gehen.

Am wichtigsten ist der *Multiplikationssatz:*

Satz: *a)* Für zwei Matrizen $A, B \in k^{n\times n}$ ist

$$\det(AB) = \det A \cdot \det B\,.$$

b) Für eine invertierbare Matrix $A \in k^{n\times n}$ ist

$$\det(A^{-1}) = (\det A)^{-1}\,.$$

Beweis: a) Die Determinante von AB ist die Determinante der Spaltenvektoren $\vec v_1, \dots, \vec v_n$ von AB, die von B ist gleich der Determinante der Spaltenvektoren $\vec b_1, \dots, \vec b_n$ von B, und nach Definition der Matrizenmultiplikation ist mit $A = (a_{ij})$

$$\vec v_i = \sum_{j=1}^{n} a_{ij} \vec b_j\,.$$

Damit sind wir in der Situation der Abschnitte *b)* und *d)* (wobei hier die $\vec b_j$ die Rolle der dortigen $\vec e_j$ übernehmen) und können aus den Eigenschaften (D1) bis (D3) folgern, daß

$$\det\big(\vec v_1, \dots, \vec v_n\big) = \left(\sum_{\pi \in \mathfrak{S}_n} \varepsilon(\pi) a_{1\pi(1)} \cdots a_{n\pi(n)}\right) \cdot \det\big(\vec b_1, \dots, \vec b_n\big)$$

ist. Das ist aber genau die Formel

$$\det(AB) = (\det A) \cdot (\det B)\,.$$

b) Für eine invertierbare Matrix A ist $A \cdot A^{-1} = E$ die Einheitsmatrix, und deren Determinante ist gleich der Determinante der Basisvektoren in ihrer natürlichen Reihenfolge, also 1. Daher ist nach Teil *a)*

$$\det A \cdot \det\left(A^{-1}\right) = \det E = 1 ,$$

woraus die Behauptung folgt. ∎

Der Multiplikationssatz läßt sich in ein Berechnungsverfahren für Determinanten übersetzen: Falls wir die LR-Zerlegung $LA = R$ der Matrix A kennen, sagt uns der Satz, daß

$$\det A = \frac{\det R}{\det L}$$

ist. Die Determinanten von L und R lassen sich leicht ausrechnen nach dem folgenden

Lemma: Die Determinante einer (unteren oder oberen) Dreiecksmatrix ist gleich dem Produkt ihrer Diagonalelemente.

Beweis: $A = (a_{ij})$ sei eine Dreiecksmatrix. Wie üblich ist

$$\det A = \sum_{\pi \in \mathfrak{S}_n} \varepsilon(\pi) a_{1\pi(1)} \cdots a_{n\pi(n)} .$$

Nun ist aber für eine untere *bzw.* obere Dreiecksmatrix $a_{ij} = 0$ falls $i < j$ *bzw.* $i > j$ ist. Da Permutationen bijektive Abbildungen sind, muß es für jede Permutation außer der Identität mindestens ein i geben mit $\pi(i) < i$ und mindestens ein j mit $\pi(j) > j$, d.h. das Produkt $a_{1\pi(1)} \cdots a_{n\pi(n)}$ enthält für jede nichtidentische Permutation mindestens einen Faktor Null. Damit kann höchstens die Identität einen von Null verschiedenen Beitrag zur Summe liefern; da sie Vorzeichen +1 hat ist also

$$\det A = a_{11} \cdots a_{nn}$$

das Produkt der Diagonalelemente von A. ∎

So wie wir die LR-Zerlegung definiert haben, ist L eine untere Dreiecksmatrix mit lauter Einsen in der Hauptdiagonale und R eine beliebige

obere Dreiecksmatrix. Also ist $\det L = 1$, die Determinante von R ist gleich dem Produkt der Diagonalelemente von R, und damit ist auch

$$\det A = \text{Produkt der Diagonalelemente von } R .$$

Zumindest für große n ermöglicht dies eine erheblich effizientere Berechnung von Determinanten als die definierende Formel: Der Aufwand für eine LR-Zerlegung ist ungefähr proportional zu n^3, was für große n weitaus günstiger ist als der überexponentiell steigende Aufwand proportional $n \cdot n!$ für die Summe aus der definierenden Formel.

Grundsätzlich läßt sich eine Determinante umso einfacher ausrechnen, je mehr Nullen sie enthält und (vor allem, wenn man ohne Hilfsmittel rechnet) je einfacher die Zahlen sind, die darin vorkommen. Im folgenden sollen daher einige Regeln zusammengestellt werden, mit denen sich Determinanten auf eine angenehmere Form bringen lassen.

Eine Determinante ändert ihren Wert nicht, wenn man ein Vielfaches einer Spalte zu einer anderen Spalte addiert: Die Determinante einer Matrix ist nach Definition die Determinante der Spaltenvektoren, und wenn man das λ-fache der j-ten Spalte zur i-ten Spalte addiert, ist

$$\begin{aligned} &\det\left(\vec{v}_1, \ldots, \vec{v}_i + \lambda\vec{v}_j, \ldots, \vec{v}_n\right) \\ =&\det\left(\vec{v}_1, \ldots, \vec{v}_i, \ldots, \vec{v}_n\right) + \lambda \det\left(\vec{v}_1, \ldots, \vec{v}_j, \ldots, \vec{v}_n\right) \\ =&\det\left(\vec{v}_1, \ldots, \vec{v}_i, \ldots, \vec{v}_n\right) , \end{aligned}$$

da der zweite Summand den Vektor $\vec{v}_j$ doppelt enthält und somit verschwindet.

Etwas entsprechendes gilt auch für Zeilen, aber bevor wir das beweisen können, sollten wir uns zunächst das Vertauschen der Zeilen und Spalten einer Matrix allgemein ansehen. Dazu führen wir zu einer Matrix A eine neue Matrix ${}^tA$ ein, deren Spalten genau die Zeilen von A sind und umgekehrt:

Definition: Die *transponierte Matrix* zur $n \times m$-Matrix $A = (a_{ij})$ ist die $m \times n$-Matrix ${}^tA = (b_{ij})$ mit $b_{ij} = a_{ji}$.

Beispielsweise ist also

$$
{}^t\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix} = \begin{pmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{pmatrix} \quad \text{und} \quad {}^t\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix} = \begin{pmatrix} 1 & 4 & 7 \\ 2 & 5 & 8 \\ 3 & 6 & 9 \end{pmatrix} .
$$

Im Fall einer $n \times n$-Matrix ist die transponierte Matrix einfach die an der Hauptdiagonalen gespiegelte Matrix, und allgemein sind die Spaltenvektoren der transponierten Matrix gerade die Zeilenvektoren der ursprünglichen Matrix und umgekehrt.

Lemma: Für jede $n \times n$-Matrix A ist $\det A = \det {}^tA$.

Beweis: Nach Definition ist für $A = (a_{ij})$

$$
\det A = \sum_{\pi \in \mathfrak{S}_n} \varepsilon(\pi) a_{1\pi(1)} \cdots a_{n\pi(n)}
$$

und

$$
\det {}^tA = \sum_{\pi \in \mathfrak{S}_n} \varepsilon(\pi) a_{\pi(1)1} \cdots a_{\pi(n)n} .
$$

Ordnet man die Faktoren des Produkts $a_{\pi(1)1} \cdots a_{\pi(n)n}$ nach dem ersten Index, so erhält man

$$
a_{\pi(1)1} \cdots a_{\pi(n)n} = a_{1\pi^{-1}(1)} \cdots a_{1\pi^{-1}(n)}
$$

und damit

$$
\det {}^tA = \sum_{\pi \in \mathfrak{S}_n} \varepsilon(\pi) a_{1\pi^{-1}(1)} \cdots a_{1\pi^{-1}(n)} .
$$

Da mit π auch π^{-1} die gesamte Gruppe $\mathfrak{S}_n$ durchläuft und $\varepsilon(\pi) = \varepsilon(\pi^{-1})$ ist, stimmt die Summe dieser Terme mit der aus der Formel für $\det A$ überein, die beiden Determinanten sind also gleich. ■

Für Leser, denen nicht klar ist, warum $\varepsilon(\pi)$ und $\varepsilon(\pi^{-1})$ gleich sind, seien zur Übung des Umgangs mit Permutationen hier kurz zwei Beweise angegeben: Am einfachsten sieht man diese Formel aufgrund des obigen Lemmas, wonach für zwei beliebige Permutationen π und π' gilt: $\varepsilon(\pi \circ \pi') = \varepsilon(\pi) \cdot \varepsilon(\pi')$. Setzt man hier $\pi' = \pi^{-1}$, so folgt

$$
1 = \varepsilon(\text{Identität}) = \varepsilon(\pi) \cdot \varepsilon(\pi^{-1}) ,
$$

so daß die beiden Vorzeichen gleich sein müssen.

Alternativ können wir uns auch explizit überlegen, daß sich π^{-1} als Produkt von r Transpositionen schreiben läßt, falls dies für π der Fall ist; genauer gilt: Ist $\pi = \tau_1 \circ \cdots \circ \tau_r$, so ist $\pi^{-1} = \tau_r \circ \cdots \circ \tau_1$, denn wie wir bereits im Abschnitt über Permutationen gesehen haben, ist

$$
\pi \circ (\tau_r \circ \cdots \circ \tau_1) = (\tau_1 \circ \cdots \circ \tau_r) \circ (\tau_r \circ \cdots \circ \tau_1) = \text{Identität} ,
$$

da jede Transposition zu sich selbst invers ist.

Aus obigem Lemma folgt insbesondere, daß man die Determinante einer Matrix auch als Determinante der Zeilenvektoren definieren kann und daß die Spaltenvektoren einer Matrix genau dann linear unabhängig sind, wenn es die Zeilenvektoren sind. Mit nur wenig mehr Aufwand könnte man sogar zeigen, daß der im Zusammenhang mit der Lösung linearer Gleichungssysteme definierte *Rang* einer Matrix genausogut über Zeilen- wie über Spaltenoperationen berechnet werden kann.

Zurück zu den Rechenregeln für Determinanten! Aus dem gerade bewiesenen Lemma folgt sofort:

Eine Determinante ändert ihren Wert nicht, wenn man ein Vielfaches einer Zeile zu einer anderen Zeile addierts, denn die Zeilen einer Matrix sind die Spalten der transponierten Matrix.

Ebenso zeigt man: *Eine Determinante wird mit -1 multipliziert, wenn man entweder zwei Zeilen oder zwei Spalten miteinander vertauscht,* denn für Spalten ist das gerade eine der definierenden Eigenschaften der Determinanten.

Als letzte Regel sei noch aufgeführt: *Eine Determinante wird mit λ multipliziert, wenn* eine *ihrer Zeilen oder* eine *ihrer Spalten mit $lambda$ multipliziert werden.* Für Spalten folgt das sofort aus der Linearität in jedem Argument, für Zeilen folgt es daraus nach Transponieren.

g) Der Entwicklungssatz von Laplace

Hier geht es um ein Berechnungsverfahren, das die definierenden Eigenschaften der Determinante ausnutzt, die „Entwicklung nach einer Zeile oder Spalte".

Für die Entwicklung einer Matrix nach beispielsweise nach der j-ten Spalte schreiben wir den j-ten Spaltenvektor

$$\vec{v}_j = \begin{pmatrix} a_{1j} \\ \vdots \\ a_{nj} \end{pmatrix}$$

als Linearkombination der Einheitsvektoren

$$\vec{e}_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad \vec{e}_2 = \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \quad \dots \quad \vec{e}_n = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix},$$

d.h. als

$$\vec{v}_j = a_{1j}\vec{e}_1 + \cdots + a_{nj}\vec{e}_n\,.$$

Dann ist wegen der Linearität der Determinante in ihrem j-ten Argument

$$\begin{aligned} \det A = \det\bigl(\vec{v}_1,\dots,\vec{v}_j,\dots,\vec{v}_n\bigr) &= \det\Bigl(\vec{v}_1,\dots,\sum_{i=1}^n a_{ij}\vec{e}_i,\dots,\vec{v}_n\Bigr) \\ &= \sum_{i=1}^n a_{ij}\det\bigl(\vec{v}_1,\dots,\vec{e}_i,\dots,\vec{v}_n\bigr)\,. \end{aligned}$$

Um daraus die Determinante von A zu berechnen, müssen wir die Determinanten

$$D_{ij} \underset{\text{def}}{=} \det\bigl(\vec{v}_1,\dots,\vec{e}_i,\dots,\vec{v}_n\bigr)$$

kennen, wobei $\vec{e}_i$ an der j-ten Stelle steht.

Dazu wenden wir die definierende Formel an: In den Produkten $a_{1\pi(1)}\cdots a_{n\pi(n)}$ ist für die gewählte Stelle j, an der wir den Vektor $\vec{e}_i$ eingefügt haben, $a_{j\ell} = 0$ für $\ell \neq i$ und $a_{ji} = 1$. Also verschwindet $a_{1\pi(1)}\cdots a_{n\pi(n)}$ für jede Permutation π mit $\pi(j) \neq i$.

Am einfachsten läßt sich dies im Fall $i = j = n$ weiter ausrechnen: Dann müssen wir nur Permutationen mit $\pi(n) = n$ betrachten, und die entsprechen eineindeutig den Permutationen auf der Menge aller Zahlen von 1 bis $n-1$; was übrigbleibt ist also (wegen $a_{nn} = 1$) gerade die Determinante jener $(n-1)\times(n-1)$-Matrix, die durch Streichung der letzten Spalte und der letzten Zeile aus A entsteht.

Der allgemeine Fall läßt sich durch Vertauschungen darauf zurückführen: Wir können zwar nicht einfach die j-te Spalte mit der n-ten vertauschen, denn dann steht ja die n-te Spalte an j-ter Stelle, aber wir können die j-te Spalte durch sukzessive Vertauschung mit ihrem rechten Nachbarn immer weiter nach außen bringen: Wir vertauschen zunächst die j-te Spalte mit der $(j+1)$-ten, dann die neue $(j+1)$-te (=alte j-te) Spalte mit der $(j+2)$-ten *usw.*, bis schließlich die alte j-te Spalte an n-ter Stelle steht, ohne daß sich vor oder nach der Stelle j irgendetwas an der relativen Reihenfolge der Spalten geändert hätte. Die Anzahl der dazu notwendigen Vertauschungen ist $(n-j)$, die Determinante wurde bei der gesamten Prozedur also mit $(-1)^{n-j}$ multipliziert.

Als nächstes müssen wir auch noch die i-te Zeile nach unten bringen; dazu verwenden wir die gerade gezeigte Formel $\det A = \det {}^tA$. Die i-te Zeile von A ist gleich der i-ten Spalte von ${}^tA$; wie wir oben gesehen haben, läßt sie sich durch $(n-i)$ Spaltenvertauschungen zur letzen Spalte machen, wobei die Determinante mit $(-1)^{n-i}$ multipliziert wird, und durch nochmaliges Transponieren erhalten wir schließlich jene Matrix, in der die i-te Zeile zur n-ten geworden ist, ohne daß sich vor oder nach der Stelle i irgendetwas an der relativen Reihenfolge der Zeilen geändert hätte.

Die Determinante der so entstehenden Matrix ist, wie wir oben beim Spezialfall $i = j = n$ gesehen haben, gleich der Determinante jener $(n-1)\times(n-1)$-Matrix A_{ij}, die durch Streichen der letzten Zeile und Spalte entsteht; bezogen auf die ursprüngliche Matrix A entsteht sie durch Streichen der i-ten Zeile und der j-ten Spalte. Also ist, wenn wir noch die Vertauschungen berücksichtigen,

$$D_{ij} = (-1)^{n-i}\cdot(-1)^{n-j}\cdot\det A_{ij} = (-1)^{i+j}\det A_{ij}\,,$$

denn $(n-i)+(n-j) = 2n-(i+j)$ hat dieselbe Parität wie $(i+j)$. Fassen wir alles zusammen, erhalten wir die Formel

$$\det A = \sum_{i=1}^n (-1)^{i+j} a_{ij}\det A_{ij}\,.$$

Die Entwicklung nach einer Zeile geht ganz entsprechend: Da die i-te Zeile von A gleich der i-ten Spalte von ${}^tA$ ist, führt obige Rechnung für die transponierte Matrix auf die Formel

$$\det A = \sum_{j=1}^{n} (-1)^{i+j} a_{ij} \det A_{ij} .$$

Damit haben wir den folgenden Satz bewiesen, der trotz seines Namens bereits LEIBNIZ bekannt war:

Entwicklungssatz von Laplace: A sei eine $n \times n$-Matrix und A_{ij} sei jene $(n-1) \times (n-1)$-Matrix, die durch Streichung der i-ten Zeile und der j-ten Spalte von A entsteht. Dann ist für jedes feste j

$$\det A = \sum_{i=1}^{n} (-1)^{i+j} a_{ij} \det A_{ij}$$

und für jedes feste i

$$\det A = \sum_{j=1}^{n} (-1)^{i+j} a_{ij} \det A_{ij} ,$$

■

Die erste Formel aus diesem Satz bezeichnet man als die *Entwicklung der Determinante nach ihrer j-ten Spalte,* die zweite (die durch Anwendung der ersten auf die transponierte Matrix entsteht) entsprechend als die *Entwicklung nach der i-ten Zeile.*

PIERRE-SIMON DE LAPLACE (1749–1827) war einer der bedeutendsten französischen Wissenschaftler seiner Zeit; berühmt sind vor allem seine Anwendungen der Analysis auf die Wahrscheinlichkeitstheorie, die Himmelsmechanik und die Potentialtheorie. Bekannt wurde er auch durch die nach ihm und KANT benannte Theorie zur Entstehung des Universums. Als politischer Opportunist kam er gut durch die Wirren seiner Zeit; er saß unter anderem in dem Kommitee, das Maßeinheiten nach dem Dezimalsystem einführte, war kurze Zeit Innenminister und wurde schließlich sogar Graf und Marquis.

BARON GOTTFRIED WILHELM VON LEIBNIZ (1646–1716) gilt als der letzte Universalgelehrte, der das gesamte Wissen seiner Zeit überblickte. In der Mathematik ist er vor allem berühmt durch die Entwicklung der Infinitesimalrechnung (bezüglich derer es einen langen Prioritätsstreit mit NEWTON gab); Bezeichnungen wie $\frac{dy}{dx}$ und $\int f(x)\,dx$ gehen auf ihn zurück. Durch seine Begründung der symbolischen Logik legte er auch einen wesentlichen Grundstein der späteren Informatik. Weitere Arbeiten befassen sich mit den Naturwissenschaften und der Technik, der Philosophie, Theologie und der Geschichte.

Als erste Anwendung wollen wir noch einmal die SARRUSsche Regel für Determinanten von 3×3-Matrizen beweisen: Entwicklung nach der ersten Spalte zeigt, daß

$$\begin{aligned}
\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} &= a_{11} \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - a_{21} \begin{vmatrix} a_{12} & a_{13} \\ a_{32} & a_{33} \end{vmatrix} + a_{31} \begin{vmatrix} a_{12} & a_{13} \\ a_{22} & a_{23} \end{vmatrix} \\
&= a_{11}(a_{22}a_{33} - a_{23}a_{32}) - a_{21}(a_{12}a_{33} - a_{13}a_{32}) \\
&\quad + a_{31}(a_{12}a_{23} - a_{13}a_{22})
\end{aligned}$$

ist, was ausmultipliziert natürlich genau die alte Formel ergibt.

Rein theoretisch ist es gleichgültig, nach welcher Zeile oder Spalte man eine Determinante entwickelt, aber man wird natürlich versuchen, eine zu finden, in der möglichst viele Nullen stehen, so daß in obiger Summe möglichst wenige Summanden übrigbleiben. In der Tat überlegt man sich leicht, daß bei der Anwendung des LAPLACEschen Entwicklungssatzes auf eine vollbesetzte Determinante wie bei der definierenden Formel rund $n \cdot n!$ Recheenoperationen erfordert und somit gegenüber der direkten Anwendung dieser Formel nur den Vorteil bietet, diese Rechenoperationen klarer zu gliedern.

Interessant zur praktischen Berechnung von Determinanten, vorzugsweise für nicht allzu große n, ist die Formel daher nur, wenn man sie mit den Rechenregeln aus dem vorigen Abschnitt kombiniert, um sich Nullen und einfache Einträge zu verschaffen. Da dieselbe Art von Operationen auch zur LR-Zerlegung führen, wird dann der Übergang zwischen

LAPLACEschem Entwicklungssatz und Berechnung via LR-Zerlegung fließend.

Betrachten wir als Beispiel die Determinante der Matrix

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

Sie hat zwar einen Eintrag Null, aber da in der dritten und vierten Spalte gleich in zwei Zeilen die Einträge entgegengesetzt gleich sind, ist es besser, zunächst die dritte Spalte durch die Summe von dritter und vierter Spalte zu ersetzen:

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

Wenn wir jetzt nach der dritten Spalte entwickeln, kommen nur zwei der Summanden wirklich vor: Der erste ($i = 1$ und $j = 3$) und der dritte ($i = j = 3$). In beiden Fällen ist $i + j$ gerade, der Vorfaktor $(-1)^{i+j}$ ist also jeweils +1. Die 3×3-Determinanten, die durch Streichung der dritten Spalte und der ersten *bzw.* dritten Zeile entstehen, können nach der Regel von SARRUS berechnet werden:

$$\begin{vmatrix} -1 & -3 & -2 \\ -4 & -2 & 0 \\ 3 & 2 & 1 \end{vmatrix} = (2 + 0 + 16) - (12 + 0 + 12) = -6$$

und

$$\begin{vmatrix} 1 & 2 & 1 \\ -1 & -3 & -2 \\ 3 & 2 & 1 \end{vmatrix} = (-3 - 12 - 2) - (-9 - 4 - 2) = -2 .$$

Also ist $\det A = (-1)(-6) + (-1)(-2) = 8$.

Auch als Hilfmittel zur Berechnung von abstrakt gegebenen Determinanten ist der LAPLACEsche Entwicklungssatz oft nützlich, da er es bei geschickter Anwendung gelegentlich erlaubt, eine Rekursionsformel zu finden und damit eine allgemeine Formel für einen gewisse Klasse von Determinanten.

Als Beispiel betrachten wir die unter anderem in der Numerik wichtige VANDERMONDEsche Determinante

$$V(a_1, \dots, a_n) = \begin{vmatrix} 1 & a_1 & a_1^2 & \dots & a_1^{n-1} \\ 1 & a_2 & a_2^2 & \dots & a_2^{n-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & a_n & a_n^2 & \dots & a_n^{n-1} \end{vmatrix} .$$

Der Franzose ALEXANDRE THÉOPHILE VANDERMONDE (1735–1796) war zunächst Musiker; erst im Alter von 35 Jahren begann er, sich für Mathematik zu interessieren und publizierte in den Jahren 1771 und 1772 vier Arbeiten über Gleichungen, Determinanten und über das Problem, einen Springer so über ein Schachbrett zu bewegen, daß er jedes Feld genau einmal betritt. Die VANDERMONDEsche Determinante ist nirgends in seinem publizierten Werk zu finden; sie wurde erst um 1935 von HENRI LEBESGUE (1875–1941) nach ihm benannt.

Für die Anwendung des LAPLACEschen Entwicklungssatzes auf diese Determinante bietet sich an, nach der ersten Spalte zu entwickeln, denn diese besteht aus lauter Einsen, und die $(n-1) \times (n-1)$-Determinanten, die nach dem Entwicklungssatz auftreten, sind im wesentlichen wieder VANDERMONDEsche Determinanten. Allerdings entstehen dabei n Summanden, und für jeden von diesen muß die ganz Prozedur wiederholt werden *usw.*– die Berechnung der Determinante auf diese Weise ist also zumindest ziemlich aufwendig.

Ein besserer Ansatz ergibt sich, wenn wir die Einsen in der ersten Spalte dadurch ausnutzen, daß wir etwa die erste Zeile von jeder anderen Zeile subtrahieren. Der Wert der Determinante ändert sich dadurch nicht, d.h.

$$V(a_1, \dots, a_n) = \begin{vmatrix} 1 & a_1 & a_1^2 & \dots & a_1^{n-1} \\ 0 & a_2 - a_1 & a_2^2 - a_1^2 & \dots & a_2^{n-1} - a_1^{n-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & a_n - a_1 & a_n^2 - a_1^2 & \dots & a_n^{n-1} - a_1^{n-1} \end{vmatrix} .$$

Wenn wir hier nach der ersten Spalte entwickeln, muß nur eine einzige $(n-1) \times (n-1)$-Determinante berücksichtigt werden, alle anderen haben den Vorfaktor null. Also ist

$$V(a_1,\dots,a_n)=\begin{vmatrix} a_2-a_1 & a_2^2-a_1^2 & \dots & a_2^{n-1}-a_1^{n-1}\\ a_3-a_1 & a_3^2-a_1^2 & \dots & a_3^{n-1}-a_1^{n-1}\\ \vdots & \vdots & \ddots & \vdots\\ a_n-a_1 & a_n^2-a_1^2 & \dots & a_n^{n-1}-a_1^{n-1}\end{vmatrix}$$

gleich jeder Determinanten, die durch Streichung der ersten Spalte und der ersten Zeile entsteht.

Hier können wir in jeder Zeile die jeweils vorne stehende Differenz ausklammern, denn genau wie

$$x^k-1=(x-1)(x^{k-1}+x^{k-2}+\cdots+x+1)$$

durch $(x-1)$ teilbar ist, ist auch

$$a_i^k-a_1^k=(a_i-a_1)(a_i^{k-1}+a_i^{k-2}a_1+a_i^{k-3}a_1^2+\cdots+a_ia_1^{k-2}+a_1^{k-1})$$

durch (a_i-a_1) teilbar; den Quotienten schreiben wir kurz als $q_{i,k-1}$:

$$q_{i,k-1}\underset{\text{def}}{=}a_i^{k-1}+a_i^{k-2}a_1+a_i^{k-3}a_1^2+\cdots+a_ia_1^{k-2}+a_1^{k-1}\,.$$

Wegen der Linearität der Determinante können wir jeden Faktor, den wir aus einer Zeile (oder Spalte) ausklammern, vor die Determinante ziehen und erhalten für $V(a_1,\dots,a_n)$ somit den Wert

$$(a_2-a_1)(a_3-a_1)\cdots(a_n-a_1)\begin{vmatrix}1 & q_{21} & q_{22} & \dots & q_{2,n-2}\\ 1 & q_{31} & q_{32} & \dots & q_{3,n-2}\\ \vdots & \vdots & \vdots & \ddots & \vdots\\ 1 & q_{n1} & q_{n2} & \dots & q_{n,n-2}\end{vmatrix}\,.$$

Die Nützlichkeit dieser Formel steht und fällt damit, daß wir die q_{ij} gut miteinander in Verbindung bringen können. Für verschiedene Indizes i haben die entsprechenden Ausdrücke offensichtlich wenig miteinander zu tun; sie enthalten nicht einmal dieselben Variablen. Schreiben wir allerdings

$$\begin{aligned}q_{ij}&=a_i^j+a_i^{j-1}a_1+a_i^{j-2}a_1^2+\cdots a_ia_1^{j-1}+a_1^j\\ &=a_i^j+a_1(a_i^{j-1}+a_i^{j-2}a_1+\cdots a_ia_1^{j-2}+a_1^{j-1})\,,\end{aligned}$$

so sehen wir, daß

$$q_{ij}=a_i^j+a_1q_{i,j-1}\quad\text{oder}\quad q_{ij}-a_1q_{i,j-1}=a_i^j$$

ist. Subtrahieren wir also zuerst a_1 mal die vorletzte Spalte von der letzten, so werden die Einträge der letzten Spalte zu a_i^{n-2}. Entsprechend subtrahieren wir a_1 mal die $(n-2)$-te Zeile von der $(n-1)$-ten und erhalten lauter Einträge a_i^{n-3} und so weiter, bis schließlich die Subtraktion des a_1-fachen der ersten Spalte von der zweiten die Einträge der letzteren zu

$$q_{i1}-a_1=(a_i+a_1)-a_1=a_i$$

macht. Somit ist $V(a_1,\dots .a_n)$ gleich

$$(a_2-a_1)(a_3-a_1)\cdots(a_n-a_1)\begin{vmatrix}1 & a_2 & a_2^2 & \dots & a_2^{n-2}\\ 1 & a_3 & a_3^2 & \dots & a_3^{n-2}\\ \vdots & \vdots & \vdots & \ddots & \vdots\\ 1 & a_n & a_n^2 & \dots & a_n^{n-2}\end{vmatrix}\,.$$

Die Determinante rechts ist offensichtlich wieder eine VANDERMONDEsche Determinante, allerdings mit um eins verminderter Zeilen- und Spaltenzahl und mit einer Variablen weniger.

Damit haben wir die Rekursionsformel

$$V(a_1,\dots,a_n)=(a_2-a_1)(a_3-a_1)\cdots(a_n-a_1)\,V(a_2,\dots,a_n)\,,$$

die es erlaubt, die Berechnung von $V(a_1,\dots,a_n)$ auf eine einzige VANDERMONDEsche Determinante der Größe $(n-1)\times(n-1)$ zurückzuführen.

Zur vollständigen Berechnung von $V(a_1,\dots,a_n)$ fehlt uns jetzt nur noch ein Induktionsanfang; direktes Nachrechnen zeigt sofort, daß

$$V(a_n)=\det(1)\quad\text{und}\quad V(a_{n-1},a_n)=\begin{vmatrix}1 & a_{n-1}\\ 1 & a_n\end{vmatrix}=a_n-a_{n-1}$$

ist, also folgt induktiv

$$V(a_1,\dots,a_n)=\prod_{j<i}(a_i-a_j)\,.$$

h) Die Cramersche Regel

Determinanten können auch angewandt werden, um die Lösungen eines linearen Gleichungssystems vom Rang n aus n Gleichungen in n Unbekannten in geschlossener Form als Funktion der Koeffizienten darzustellen. Dazu schreiben wir das Gleichungssystem

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

in der Form

$$x_1\vec{a}_1 + \cdots + x_n\vec{a}_n = \vec{b}\,,$$

wobei die $\vec{a}_i$ die Spaltenvektoren der Matrix A seien und $(x_1, \ldots, x_n)$ eine Lösung des linearen Gleichungssystems.

Nun ersetzen wir in

$$\det A = \det\bigl(\vec{a}_1, \ldots, \vec{a}_n\bigr)$$

rechts den Vektor $\vec{a}_i$ durch die rechte Seite $\vec{b}$ des Gleichungssystems. (Es gibt eigentlich keinen vernünftigen Grund, warum wir das tun sollten; auch dieser Trick wird, wie so viele, erst nachträglich durch das Ergebnis gerechtfertigt.) Die so entstehende Determinante ist

$$\begin{aligned}
&\det\bigl(\vec{a}_1, \ldots, \vec{b}, \ldots, \vec{a}_n\bigr)\\
&= \det\Bigl(\vec{a}_1, \ldots, \sum_{j=1}^{n} x_j\vec{a}_j, \ldots, \vec{a}_n\Bigr)\\
&= \sum_{j=1}^{n} x_j \det\bigl(\vec{a}_1, \ldots, \vec{a}_j, \ldots, \vec{a}_n\bigr)\\
&= x_i \det\bigl(\vec{a}_1, \ldots, \vec{a}_i, \ldots, \vec{a}_n\bigr) = x_i \det A\,,
\end{aligned}$$

da für jeden Index $j \neq i$ der Vektor $\vec{a}_j$ zweimal als Argument der Determinante auftritt, so daß alle Summanden bis auf den i-ten verschwinden.

Falls $\det A = 0$ ist, nützt uns diese Formel überhaupt nichts; ist allerdings $\det A \neq 0$, wissen wir bereits, daß das Gleichungssystem eindeutig lösbar ist, und wir können die Komponenten dieser eindeutig bestimmten Lösung explizit ausdrücken durch die

CRAMERsche Regel:

$$x_i = \frac{\det\bigl(\vec{a}_1, \ldots, \vec{b}, \ldots, \vec{a}_n\bigr)}{\det A}$$

Der Schweizer Mathematiker GABRIEL CRAMER (1704–1752) lehrte an der Universität Genf. Bekannt wurde er vor allem durch seine Arbeiten über Determinanten, er beschäftigte sich aber auch viel mit Analysis und Geometrie, insbesondere ist er Autor eines Buchs über algebraische Kurven. Weitere Arbeitsgebiete sind mathematische Methoden der Physik sowie die Geschichte der Mathematik. Die CRAMERsche Regel im Spezialfall $n = 2$ ist bereits 1545 in der *Ars magna* des italienischen Mathematikers GIROLAMO CARDANO (1501–1576) zu finden.

Die CRAMERsche Regel ist sicherlich kein Verfahren, das man oft anwendet zur Lösung eines einzelnen Gleichungssystems: Der Aufwand für die Berechnung von $n + 1$ Determinanten ist weit größer als eine Lösung nach dem GAUSS-Algorithmus. Falls man allerdings eine ganze Familie ähnlicher Gleichungssysteme hat, in der sich nur wenige Parameter ändern, kann es sich lohnen, mit Hilfe der CRAMERschen Regel eine (von den Parametern abhängige) Lösungsformel zu berechnen und dann diese anzuwenden.