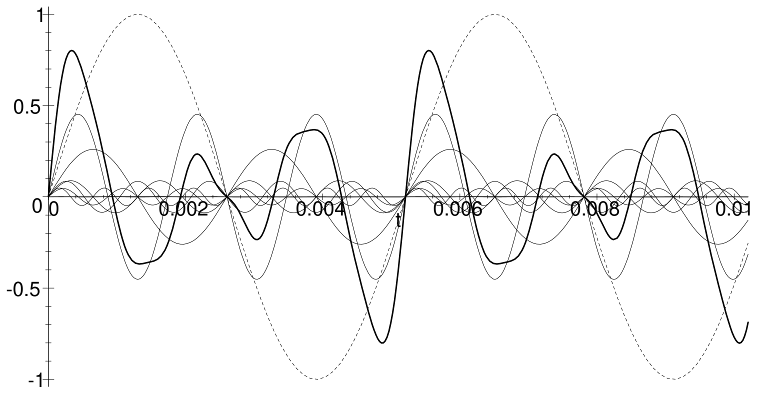

Abb. 9: Ton der g-Saite einer Geige und seine Komponenten

Linearkombinationen sind also ein einfaches Mittel, um aus relativ wenigen einfachen Funktionen oder Vektoren kompliziertere aufzubauen. Am besten geht das, wenn man mit möglichst wenig Bausteinen auskommt, und das ist genau dann der Fall, wenn die Menge M, aus deren Elementen wir Linearkombinationen bilden, linear unabhängig ist. Für eine linear abhängige Menge können wir nämlich nach dem folgenden Lemma mindestens ein Element eliminieren:

Lemma: Falls die Menge $M \subseteq V$ linear abhängig ist, gibt es ein Element $\vec{v} \in M$, das sich als Linearkombination der übrigen, d.h. von Vektoren aus $M \setminus \{\vec{v}\}$ schreiben läßt. Insbesondere ist dann auch

$$[M] = \left[M \setminus \{\vec{v}\}\right] .$$

Beweis: Wenn M linear abhängig ist, gibt es eine nichttriviale Linearkombination von Vektoren $\vec{v}_i \in M$, so daß

$$\lambda_1\vec{v}_1 + \cdots + \lambda_n\vec{v}_n = \vec{0}$$

ist mit Körperelementen λ_i, die nicht alle gleich Null sind. Sei zum Beispiel $\vec{\lambda}_j \neq 0$. Dann kann obige Gleichung nach $\vec{v}_j$ aufgelöst werden; für $1 < j < n$ etwa ist

$$\vec{v}_j = -\frac{\lambda_1}{\lambda_j}\vec{v}_1 - \cdots - \frac{\lambda_{j-1}}{\lambda_j}\vec{v}_{j-1} - \frac{\lambda_{j+1}}{\lambda_j}\vec{v}_{j+1} - \cdots - \frac{\lambda_n}{\lambda_j}\vec{v}_n ,$$

und entsprechend läßt sich $\vec{v}_j$ auch im Falle $j = 1$ oder $j = n$ als Linearkombination der übrigen $\vec{v}_i$ schreiben.

■

g) Die Dimension eines Vektorraums

Die Dimension eines Vektorraums soll natürlich so definiert werden, daß $\mathbb{R}^n$ die Dimension n hat; wir müssen die Zahl n also irgendwie als Eigenschaft von (Mengen von) Vektoren aus $\mathbb{R}^n$ rekonstruieren.

Offensichtlich kann jeder Vektor aus $\mathbb{R}^n$ als Linearkombination der n Einheitsvektoren geschrieben werden:

$$\begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} = a_1 \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + a_2 \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix} + \cdots + a_n \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix} .$$

Es fällt aber schwer sich eine Menge aus weniger als n Vektoren vorzustellen, aus der sich ebenfalls *jeder* Vektor aus $\mathbb{R}^n$ als Linearkombination darstellen läßt.

Diese Eigenschaft machen wir uns zunutze, um allgemein Dimensionen zu definieren:

Definition: *a)* Wir sagen, der k-Vektorraum V sei *endlichdimensional,* wenn es eine endliche Teilmenge $M \subseteq V$ gibt, so daß $[M] = V$ ist; ansonsten bezeichen wir V als *unendlichdimensional.*
b) Wir sagen, der endlichdimensionale Vektorraum V habe die Dimension n, in Zeichen $n = \dim_k V$ oder kurz $n = \dim V$, wenn er eine n-elementige Teilmenge $M \subseteq V$ enthält mit $[M] = V$, aber keine Menge M' mit weniger als n Elementen, für die $[M'] = V$ ist.
c) Dem Nullvektorraum $\{\vec{0}\}$ ordnen wir (formal) die Dimension null zu.

Als Beispiel eines unendlichdimensionalen Vektorraums haben wir den Vektorraum aller reeller Polynome. Hätte dieser nämlich ein endliches Erzeugendensystem, bestehend etwa aus den Polynomen f_1 bis f_n, so ließe sich sich jedes Polynom als Linearkombination

$$f = \lambda_1 f_1 + \cdots + \lambda_n f_n$$

schreiben. Auf diese Weise aber erhält man nur Polynome, deren Grad nicht größer ist als der größte Grad eines f_i.

Endlichdimensionale Vektorräume sind insbesondere die reellen Vektorräume $\mathbb{R}^n$, denn $\mathbb{R}^n$ wird von seinen n Einheitsvektoren erzeugt. Da wir aber noch nicht sicher wissen, daß es kein Erzeugendensystem mit *weniger* als n Vektoren gibt, können wir im Augenblick nur sagen, daß die Dimension von $\mathbb{R}^n$ *höchstens* n ist.

Für $\mathbb{R}^2$ sieht man leicht, daß sie genau zwei ist: Ansonsten gäbe es nämlich ein Erzeugendensystem aus nur einem Vektor, d.h. alle Vektoren aus $\mathbb{R}^2$ wären proportional zueinander, was natürlich nicht der Fall ist. Für beliebiges n müssen wir ähnlich argumentieren mit linearer Abhängigkeit anstelle von Proportionalität; Die Methoden dazu entwickelt der nächste Abschnitt.

h) Basen

Im $\mathbb{R}^3$ läßt sich jeder Vektor

$$\vec{v} = \begin{pmatrix} a \\ b \\ c \end{pmatrix} = a \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + b \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} + c \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

auf genau eine Weise als Linearkombination der drei Einheitsvektoren

$$\vec{e}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad \vec{e}_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \quad \text{und} \quad \vec{e}_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

schreiben; die entspricht der Tatsache, daß wir im $\mathbb{R}^3$ drei Koordinaten haben.

Mit dem Begriff der *Basis* soll dieser Sachverhalt (soweit möglich) auf beliebige Vektorräume verallgemeinert werden.

Definition: Eine Teilmenge $\mathcal{B} \subset V$ eines k-Vektorraums V heißt *Basis* von V, wenn gilt:
1.) $V = [\mathcal{B}]$, d.h. V wird von $\mathcal{B}$ erzeugt, *und*
2.) $\mathcal{B}$ ist linear unabhängig.

In diesem Sinne ist $\{\vec{e}_1, \vec{e}_2, \vec{e}_3\}$ offenbar eine Basis des $\mathbb{R}^3$. Ihre wesentliche Eigenschaft der eindeutigen Darstellbarkeit eines jeden Vektors als Linearkombination teilt sie mit jeder anderen Basis:

Lemma: Ist $\mathcal{B}$ eine Basis eines Vektorraums V, so läßt sich jeder Vektor $\vec{v} \in V$ auf genau eine Weise als Linearkombination

$$\vec{v} = \lambda_1 \vec{b}_1 + \cdots + \lambda_r \vec{b}_r$$

von Basisvektoren $\vec{b}_i \in \mathcal{B}$ darstellen.

Beweis: Nach der ersten Eigenschaft in der Definition einer Basis ist $V = [\mathcal{B}]$, also läßt sich jeder Vektor $\vec{v} \in V$ als Linearkombination von endlich vielen Elementen aus $\mathcal{B}$ darstellen. Auch wenn wir von zwei solchen Darstellungen ausgehen, ist die Menge der daran beteiligten Vektoren aus $\mathcal{B}$ noch endlich; wir können also annehmen, daß es r Vektoren $\vec{b}_1, \ldots, \vec{b}_r$ gibt, so daß

$$\begin{aligned} \vec{v} &= \lambda_1 \vec{b}_1 + \cdots + \lambda_r \vec{b}_r \\ &= \mu_1 \vec{b}_1 + \cdots + \mu_r \vec{b}_r \end{aligned}$$

ist, wobei wir einfach λ_i oder μ_i auf Null setzen, wenn $\vec{b}_i$ in der entsprechenden Darstellung nicht vorkommt. Subtrahieren wir die beiden Darstellungen voneinander, erhalten wir eine Darstellung des Nullvektors als Linearkombination

$$\vec{0} = (\lambda_1 - \mu_1)\vec{b}_1 + \cdots + (\lambda_r - \mu_r)\vec{b}_r$$

von Vektoren aus der Menge $\mathcal{B}$. Da diese nach der zweiten definierenden Eigenschaft einer Basis linear unabhängig ist, müssen alle Koeffizienten $\lambda_i - \mu_i$ verschwinden. Damit sind die beiden betrachteten Darstellungen von $\vec{v}$ als Linearkombination der Vektoren aus $\mathcal{B}$ gleich, mit anderen Worten: Es gibt genau eine solche Darstellung. ■

Lemma: Ein Erzeugendensystem ist genau dann eine Basis, wenn es minimal ist.

Beweis: Das Erzeugendensystem $\mathcal{B}$ sei eine Basis. Um zu zeigen, daß es minimal ist, müssen wir uns überlegen, daß $\mathcal{B} \setminus \{\vec{v}\}$ für keinen der Vektoren $\vec{v} \in \mathcal{B}$ ein Erzeugendensystem ist.

Wäre $\mathcal{B} \setminus \{\vec{v}\}$ ein Erzeugendensystem, so könnte insbesondere der Vektor $\vec{v}$ als Linearkombination der Vektoren aus $\mathcal{B} \setminus \{\vec{v}\}$ geschrieben werden. Gleichzeitig hat er aber die Darstellung $\vec{v} = \vec{v}$, deren rechte Seite man auch als Linearkombination von Elementen aus $\mathcal{B}$ auffassen kann. Somit ist seine Basisdarstellung nicht eindeutig, im Widerspruch zum gerade bewiesenen Lemma. Daher muß $\mathcal{B}$ minimal sein.

Umgekehrt sei $\mathcal{B}$ ein minimales Erzeugendensystem. Um zu zeigen, daß es eine Basis ist, reicht der Nachweis der linearen Unabhängigkeit von $\mathcal{B}$.

Sei also $\lambda_1\vec{b}_1 + \cdots + \lambda_n\vec{b}_n = \vec{0}$ eine Darstellung des Nullvektors als Linearkombination von Elementen aus $\mathcal{B}$. Falls darin einer der Koeffizienten λ_i nicht verschwindet, läßt sich der zugehörige Vektor $\vec{v}_i$ als Linearkombination der restlichen $\vec{v}_j$ schreiben. Dann reicht aber bereits $\mathcal{B} \setminus \{\vec{v}_i\}$ zur Erzeugung aus, $\mathcal{B}$ ist also nicht minimal. Somit müssen alle λ_i verschwinden, $\mathcal{B}$ ist also linear unabhängig und damit eine Basis. ■

Lemma: Eine linear unabhängige Teilmenge eines Vektorraums ist genau dann eine Basis, wenn sie maximal ist.

Beweis: Die linear unabhängige Menge $\mathcal{B}$ sei eine Basis. Dann läst sich jeder Vektor $\vec{v} \notin \mathcal{B}$ als Linearkombination der Elemente von $\mathcal{B}$ schreiben, $\mathcal{B} \cup \{\vec{v}\}$ ist also nicht linear abhängig.

Umgekehrt sei $\mathcal{B}$ eine maximale linear unabhängige Teilmenge und $\vec{v}$ sei ein beliebiger Vektor. Falls $\vec{v}$ in $\mathcal{B}$ liegt, liegt er erst recht im Erzeugnis $[\mathcal{B}]$ von $\mathcal{B}$, andernfalls ist $\mathcal{B} \cup \{\vec{v}\}$ linear abhängig, es gibt also eine nichttriviale Linearkombination

$$\lambda\vec{v} + \lambda_1\vec{b}_1 + \cdot + \lambda_n\vec{b}_n = \vec{0}$$

mit Vektoren $\vec{b}_i \in \mathcal{B}$. Darin muß $\lambda \neq 0$ sein, denn sonst wäre $\mathcal{B}$ linear abhängig. Also liegt

$$\vec{v} = -\frac{\lambda_1}{\lambda}\vec{b}_1 + \cdots + -\frac{\lambda_n}{\lambda}\vec{b}_n$$

in $[\mathcal{B}]$, und $\mathcal{B}$ ist ein Erzeugendensystem. ■

Basen lassen sich also auch charakterisieren als minimale Erzeugendensysteme oder als maximale linear unabhängige Teilmengen eines Vektorraums.

Als nächstes stellt sich die Frage, wann es Basen gibt. Glücklicherweise hat *jeder* Vektorraum eine Basis, der Beweis ist allerdings für unendlichdimensionale Vektorräume logisch nicht ganz einfach. Für diese Vorlesung wollen wir uns daher mit einem Beweis für endlichdimensionale Vektorräume begnügen. Wir beweisen dazu den etwas allgemeineren, tatsächlich ebenfalls für beliebige Vektorräume gültigen

Basisergänzungssatz: $M \subset V$ sei eine linear unabhängige Teilmenge des endlichdimensionalen Vektorraums V. Dann gibt es eine Basis $\mathcal{B}$ von V, die M enthält.

Beweis: Da V nach Voraussetzung endlichdimensional ist, gibt es zunächst einmal überhaupt eine endliche Menge $E \subset V$, die V erzeugt. Falls E einen oder mehrere der Vektoren aus M enthält, entfernen wir diese; was übrigbleibt, sei die Menge N, d.h. $N = E \setminus M$.

Konkret sei $M = \{\vec{b}_1, \ldots, \vec{b}_r\}$ und $N = \{\vec{v}_1, \ldots, \vec{v}_s\}$; dann wird V auch erzeugt von

$$M \cup N = \{\vec{b}_1, \ldots, \vec{b}_r, \vec{v}_1, \ldots, \vec{v}_s\} .$$

Wir beweisen die Behauptung durch Induktion nach s.

Für $s = 0$ ist M bereits eine Basis, und wir sind fertig.

Für $s > 0$ sind wir fertig, falls $M \cup N$ linear unabhängig ist; denn dann ist $M \cup N$ eine Basis von V, die M enthält.

Andernfalls gibt es Elemente $\lambda_1, \ldots, \lambda_r$ und $\mu_1, \ldots, \mu_s$, die nicht alle gleichzeitig verschwinden, so daß

$$\lambda_1\vec{b}_1 + \cdots + \lambda_r\vec{b}_r + \mu_1\vec{v}_1 + \cdots + \mu_s\vec{v}_s = \vec{0}$$

ist. In dieser Gleichung können nicht alle μ_i verschwinden, denn sonst wären die $\vec{b}_i$ linear abhängig, im Widerspruch zur Voraussetzung. Also gibt es (mindestens) ein $\mu_i \neq 0$, und der zugehörige Vektor $\vec{v}_i$ läßt sich als Linearkombination der restlichen $\vec{v}_j$ und der $\vec{b}_\ell$ ausdrücken.

Damit wird $V = [M \cup (N \setminus \{\vec{v}_i\})]$ auch von der um $\vec{v}_i$ verminderten Menge erzeugt, und wir haben nur noch $s-1$ Vektoren $\vec{v}_j$. Daher gibt es nach Induktionsannahme eine Basis von V, die M enthält. ■

Korollar: Jeder endlichdimensionale Vektorraum $V \neq \{\vec{0}\}$ hat eine Basis.

Beweis: Man wende den obigen Satz an auf eine Menge M, die aus einem einzigen Vektor $\vec{v} \neq \vec{0}$ besteht. ■

Um die Sonderrolle des Nullvektorraums zu eliminieren, vereinbaren wir, daß er die leere Menge als Basis haben soll; dies ist kompatibel mit der üblichen Interpretation von leeren Summen.

Wie bereits erwähnt, gelten sowohl der Basisergänzungssatz als auch das obige Korollar für beliebige Vektorräume, d.h. also auch im Falle unendlicher Dimension. Für interessierte Leser sei kurz erwähnt, wie man hier vorgeht. Das wesentliche neue Hilfsmittel ist das ZORNsche Lemma, benannt nach dem deutschen Mathematiker MAX ZORN (1906–1993), der es, nachdem er Deutschland wegen der nationalsozialistischen Politik verlassen mußte, um 1935 an der amerikanischen Yale Universität bewies. Es besagt folgendes:

Gegeben sei eine nichtleere partiell geordnete Menge $\mathcal{M}$, d.h. für manche Paare von Elementen $A, B \in \mathcal{M}$ ist eine Relation $A < B$ erklärt mit der Eigenschaft, daß mit $A < B$ und $B < C$ auch $A < C$ gilt, wohingegen nie $A < A$ ist. Diese partiell geordnete Menge habe die zusätzliche Eigenschaft, daß es zu jeder Kette

$$A_1 < A_2 < A_3 < \cdots$$

von Elementen aus $\mathcal{M}$ ein Element A_∞ gebe mit der Eigenschaft, daß $A_i < A_\infty$ für alle i. Dann gibt es in $\mathcal{M}$ ein *maximales* Element, d.h. ein Element B, zu dem es kein $C \in \mathcal{M}$ gibt mit $B < C$.

Dieses Lemma kann nicht aus den üblichen Axiomen der Mengenlehre hergeleitet werden, sondern ist äquivalent zum sogenannten *Auswahlaxiom.* Für dieses bewies um 1940 der österreichische Mathematiker KURT GÖDEL (1906–1978), seit 1940 im amerikanischen Exil in Princeton, daß sowohl dieses Axiom als auch seine Negation mit den restlichen Axiomen der Mengenlehre kompatibel ist; das gleiche gilt demnach auch für das ZORNsche Lemma. Man kann daher wählen, ob man eine Mathematik mit oder ohne ZORNsches Lemma bevorzugt. Die meisten Mathematiker haben sich für „mit" entschieden, es gibt aber auch welche, die das ZORNsche Lemma ablehnen.

Aus dem ZORNschen Lemma folgt der Basisergänzungssatz recht einfach: Als Menge $\mathcal{M}$ nehmen wir die Menge aller linear unabhängiger Teilmengen $A \subset V$, die M enthalten; die partielle Ordnungsrelation sei die gewöhnliche (echte) Teilmengenbeziehung. Die Kettenbedingung des ZORNschen Lemmas ist offensichtlich erfüllt, denn für eine Kette

$$M \subset A_1 \subset A_2 \subset A_3 \subset \cdots$$

aus linear unabhängigen Mengen A_i, die M enthalten, ist auch

$$A_\infty = \bigcup_{i \geq 1} A_i$$

eine linear unabhängige Teilmenge von V, die M enthält, da jede endliche Menge von Vektoren aus A_∞ bereits in einer der Mengen A_m liegt. Also gibt es nach dem ZORNschen Lemma ein maximales Element $\mathcal{B} \in \mathcal{M}$. Diese Menge $\mathcal{B}$ ist linear unabhängig, da sie in $\mathcal{M}$ liegt, und sie ist eine Basis, denn gäbe es einen Vektor $\vec{v} \notin [\mathcal{B}]$, so wäre auch die Menge $\mathcal{C} = \mathcal{B} \cup \{\vec{v}\}$ linear unabhängig, im Widerspruch zur Maximalität von $\mathcal{B}$. Damit ist der Basisergänzungssatz bewiesen, und das Korollar folgt wie oben.

Um wenigstens anhand eines Beispiels zu sehen, daß auch unendlichdimensionale Vektorräume Basen haben, betrachten wir den Vektorraum V aller Polynome mit reellen Koeffizienten. Da sich ein Polynom P vom Grad d als

$$P = a_0 + a_1X + a_2X^2 + \cdots + a_dX^d$$

schreiben läßt, erzeugt die Menge

$$\mathcal{B} = \{1, X, X^2, X^3, \ldots\}$$

diesen Vektorraum. Jede Linearkombination des Nullvektors, des Polynoms $P \equiv 0$ also, aus Elementen von $\mathcal{B}$ wäre ein Polynom

$$\lambda_0 + \lambda_1X + \cdots + \lambda_nX^n\,,$$

dessen Koeffizienten zumindest teilweise von Null verschieden sind, während es selbst identisch Null wäre. Da es kein solches Polynom gibt, ist die Menge $\mathcal{B}$ linear unabhängig, also eine Basis von V.

Die Schwierigkeiten, die bei unendlichdimensionalen Vektorräumen auftreten können, sieht man, wenn man in diesem Beispiel die Polynome durch Potenzreihen (egal ob formal oder konvergent) ersetzt: Da Potenzreihen *unendliche* Summen sind, während bei Linearkombinationen nur *endliche* Summen erlaubt sind, reichen nun die X-Potenzen nicht mehr aus zur Erzeugung. Nach dem Basisergänzungssatz, der, auch wenn wir das nicht bewiesen haben, auch für unendlichdimensionale Vektorräume gilt, gibt es eine Menge von Potenzreihen, die zusammen mit der obigen Menge $\mathcal{B}$ eine Basis bilden; explizit angeben konnte diese Menge aber noch niemand, genausowenig wie eine explizite Basis für einen der Räume $\mathcal{C}^n(\mathbb{R},\mathbb{R})$.

Kehren wir also zurück zum überschaubareren endlichdimensionalen Fall, und beweisen wir dort zunächst die anschaulich fast selbstverständliche Aussage, daß jede Basis eines n-dimensionalen Vektorraums aus n Vektoren besteht. Dazu benötigen wir eine leichte Verschärfung des Basisergänzungssatzes:

Austauschsatz von STEINITZ: M sei eine endliche linear unabhängige Teilmenge des endlichdimensionalen Vektorraums V, und $\mathcal{B}$ sei eine endliche Basis von V. Dann gibt es eine Teilmenge $\mathcal{B}'$ von $\mathcal{B}$, so daß $M \cup \mathcal{B}'$ eine Basis von V ist. Diese hat genauso viele Elemente wie $\mathcal{B}$.

Mit anderen Worten: Man kann Vektoren aus $\mathcal{B}$ finden, die sich Stück für Stück gegen die Vektoren aus M *austauschen* lassen.

Der *Beweis* ist dem des Basisergänzungssatzes sehr ähnlich; mit Rücksicht auf die Anzahlaussage führen wir ihn aber durch Induktion nach der Elementanzahl m von M.

Für $m = 0$ ist $M = \emptyset$ und die Behauptung ist trivial: Wir setzen einfach $\mathcal{B}' = \mathcal{B}$.

Für $m \geq 1$ entfernen wir einen Vektor $\vec{v}$ aus M und wenden den Satz auf die Menge $M' = M \smallsetminus \{\vec{v}\}$ an. Für diese gilt er nach Induktionsannahme, es gibt also eine Teilmenge $\mathcal{C}'$ von $\mathcal{B}$, so daß $\mathcal{C} = M' \cup \mathcal{C}'$ eine Basis von V ist mit gleicher Elementanzahl wie $\mathcal{B}$. Mit dieser Basis schreiben wir

$$\vec{v} = \lambda_1\vec{v}_1 + \cdots + \lambda_{m-1}\vec{v}_{m-1} + \mu_1\vec{c}_1 + \cdots + \mu_r\vec{c}_r\,,$$

wobei $M' = \{\vec{v}_1,\ldots,\vec{v}_{m-1}\}$ und $\mathcal{C}' = \{\vec{c}_1,\ldots,\vec{c}_r\}$ sein soll.

Da $M = M' \cup \{\vec{v}\}$ linear unabhängig ist, muß in dieser Darstellung mindestens ein $\vec{\mu}_i$ von null verschieden sein. Daher läßt sich der zugehörige Vektor $\vec{c}_i$ als Linearkombination aus den restlichen $\vec{c}_j$, den $\vec{v}_\ell$ und dem Vektor $\vec{v}$ schreiben, d.h. auch

$$M' \cup \bigl(\mathcal{B} \smallsetminus \{\vec{c}_i\}\bigr) \cup \{\vec{v}\} = M \cup \bigl(\mathcal{B} \smallsetminus \{\vec{c}_i\}\bigr)$$

erzeugt ganz V. Diese Menge ist auch linear unabhängig und somit eine Basis, denn ist

$$\alpha\vec{v} + \sum_{\ell=1}^{m-1}\alpha_\ell\vec{v}_\ell + \sum_{\substack{j=1\\ j\neq i}}^{n}\beta_j\vec{b}_j = \vec{0}\,,$$

so muß zunächst α verschwinden, da $\vec{v}$ sonst als Linearkombination der $\vec{v} \in M'$ und der $\vec{c}_j$ mit $j \neq i$ dargestellt werden könnte, was wir oben durch die Wahl eines i mit $\mu_i \neq 0$ ausgeschlossen haben. Also steht hier nur eine Linearkombination von Elementen einer Basis, so daß alle α_ℓ und β_j verschwinden müssen. Mit

$$\mathcal{B}' = \bigl(\mathcal{B} \smallsetminus \{\vec{c}_i\}\bigr)$$

ist die Behauptung des Satzes also erfüllt. ■

ERNST STEINITZ (1871–1928) wurde in Schlesien geboren und studierte ab 1890 an den Universitäten Breslau und Berlin. 1894 promovierte er in Breslau, ein Jahr später wurde er Privatdozent an der Technischen Hochschule Berlin-Charlottenburg. 1910 wurde er Professor in Breslau, 1920 in der Universität Kiel. In seinem Buch *Algebraische Theorie der Körper* gab er 1910 die erste Definition eines Körpers und bewies viele Sätze, die noch heute zum Standardstoff jeder Algebra-Vorlesung gehören. Auch die Konstruktion der rationalen Zahlen als Äquivalenzklassen von Paaren ganzer Zahlen geht auf ihn zurück.

Aus dem STEINITZschen Austauschsatz folgt

Satz: *a)* Jede Basis $\mathcal{B}$ eines n-dimensionalen Vektorraums V besteht aus n Vektoren.
b) Jede Teilmenge von V mit mehr als n Elementen ist linear abhängig.
c) Keine Teilmenge von V mit weniger als n Elementen ist ein Erzeugendensystem.

Beweis: a) Da V die Dimension n hat, gibt es ein Erzeugendensystem $M = \{\vec{v}_1,\ldots,\vec{v}_n\}$ mit n-Elementen, aber keines mit weniger Elementen. Also ist M ein minimales Erzeugendensystem und somit eine Basis.

Nun sei $\mathcal{B} = \{\vec{b}_1, \ldots, \vec{b}_m\}$ irgendeine andere Basis von V. Nach dem Austauschsatz läßt sich M zu einer Basis von V ergänzen, die genauso viele Elemente hat wie $\mathcal{B}$. Da es keine Basis geben kann, die M echt enthält, muß M genauso viele Elemente enthalten wie $\mathcal{B}$, also n.

b) Jede linear unabhängige Teilmenge läßt sich zu einer Basis ergänzen, und jede Basis besteht aus n Vektoren. Also kann eine linear unabhängige Menge höchstens n Vektoren enthalten.

c) Das ist die Definition der Dimension. ■

Nach diesem Satz läßt sich die Dimension eines Vektorraums einfach dadurch bestimmen, daß man eine Basis findet und deren Elemente zählt. Insbesondere hat $\mathbb{R}^n$ als $\mathbb{R}$-Vektorraum die Dimension n, da die n Einheitsvektoren eine Basis bilden.

Weniger offensichtlich ist, daß $\mathbb{R}$ als $\mathbb{Q}$-Vektorraum unendlichdimensional ist: Dazu betrachten wir die unendliche Menge M aller Logarithmen $\ln p$ der Primzahlen. Wäre diese Menge linear abhängig, gäbe es eine nichttriviale Linearkombination

$$\lambda_1 \ln p_1 + \cdots + \lambda_r \ln p_r = 0$$

mit $\lambda_i \in \mathbb{Q}$. Multipliziert man diese Gleichung mit dem Hauptnenner der λ_i, so erhält man eine entsprechende Gleichung mit Koeffizienten $\mu_i \in \mathbb{Z}$. Dann ist

$$\mu_1 \ln p_1 + \cdots + \mu_r \ln p_r = \ln(p_1^{\mu_1} \cdot \cdots \cdot p_r^{\mu_r}) = 0$$

gleichbedeutend mit

$$p_1^{\mu_1} \cdot \cdots \cdot p_r^{\mu_r} = 1 ,$$

was wegen der Eindeutigkeit der Primzerlegung in $\mathbb{Z}$ nur gelten kann, wenn alle μ_i und damit auch alle λ_i verschwinden.

Also ist $\mathbb{R}$ als $\mathbb{Q}$-Vektorraum unendlichdimensional, und dies erklärt, warum Computer so große Schwierigkeiten mit reellen Zahlen haben: Exakt rechnen kann ein Computer nur in Teilkörpern von $\mathbb{R}$, die endlichdimensionale $\mathbb{Q}$-Vektorräume sind.

i) Lineare Abbildungen und Untervektorräume

Als nächstes wollen wir uns mit Dimensionen von Untervektorräumen, insbesondere auch Kernen und Bildern beschäftigen. Anschaulich klar und auch recht einfach zu beweisen ist der folgende

Satz: Für einen echten Untervektorraum $U < V$ eines endlichdimensionalen Vektorraums V ist $\dim U < \dim V$.

Beweis: Eine Basis von U ist auch in V linear unabhängig, läßt sich also ergänzen zu einer Basis von V. Da die Dimension eines Vektorraums gleich der Elementanzahl einer beliebigen Basis ist, folgt sofort, daß $\dim U \leq \dim V$ sein muß, und wenn beide gleich sind, ist $U = V$. ■

Interessanter ist

Satz: Für einen endlichdimensionalen Vektorraum V und eine lineare Abbildung $\varphi\colon V \to W$ ist

$$\dim \operatorname{Bild} \varphi = \dim V - \dim \operatorname{Kern} \varphi .$$

Beweis: $\vec{b}_1, \ldots, \vec{b}_r$ sei eine Basis von $\operatorname{Kern}\varphi$; falls φ injektiv ist, setzen wir $r = 0$. Nach dem Basisergänzungssatz oder (falls $r = 0$) wegen der Existenz von Basen lassen sich dann $n - r$ Vektoren $\vec{b}_{r+1}, \ldots, \vec{b}_n$ finden mit $n = \dim V$, so daß $\vec{b}_1, \ldots, \vec{b}_n$ eine Basis von V ist.

Das Bild eines beliebigen Vektors $\vec{v} = \lambda_1\vec{b}_1 + \cdots + \lambda_n\vec{b}_n$ ist dann

$$\varphi(\vec{v}) = \lambda_1\varphi(\vec{b}_1) + \cdots + \lambda_n\varphi(\vec{b}_n) = \lambda_{r+1}\varphi(\vec{b}_{r+1}) + \cdots + \lambda_n\varphi(\vec{b}_n) ,$$

da $\vec{b}_1 \ldots, \vec{b}_r$ ja auf den Nullvektor abgebildet werden. Also wird $\operatorname{Bild}\varphi$ von den Vektoren $\varphi(\vec{b}_{r+1}), \ldots, \varphi(\vec{b}_n)$ erzeugt.

Diese Vektoren sind auch linear unabhängig in W, denn ist

$$\lambda_{r+1}\varphi(\vec{b}_{r+1}) + \cdots + \lambda_n\varphi(\vec{b}_n) = \varphi(\lambda_{r+1}\vec{b}_{r+1} + \cdots + \lambda_n\vec{b}_n) = \vec{0} ,$$

so liegt $\lambda_{r+1}\vec{b}_{r+1} + \cdots + \lambda_n\vec{b}_n$ im Kern von φ.

Im Fall einer injektiven Abbildung ist $\lambda_{r+1}\vec{b}_{r+1}+\cdots+\lambda_n\vec{b}_n$ daher gleich dem Nullvektor, und damit müssen alle $\lambda_i = 0$ sein, denn die $\vec{b}_i$ sind als Basisvektoren insbesondere linear unabhängig.

Falls φ nicht injektiv ist, können wir nur sagen, daß der Vektor $\lambda_{r+1}\vec{b}_{r+1} + \cdots + \lambda_n\vec{b}_n$ im Kern von φ liegt; er ist also darstellbar als Linearkombination der Basisvektoren $\vec{b}_1, \dots, \vec{b}_r$ des Kerns:

$$\lambda_{r+1}\vec{b}_{r+1} + \cdots + \lambda_n\vec{b}_n = \lambda_1\vec{b}_1 + \cdots + \lambda_r\vec{b}_r\,.$$

Auch daraus folgt wegen der linearen Unabhängigkeit der $\vec{b}_i$, daß alle λ_i Null sein müssen.

Damit ist $\{\vec{b}_{r+1}, \dots, \vec{b}_n\}$ eine Basis von Bild φ, d.h.

$$\dim \operatorname{Bild}\varphi = n - r = \dim V - \dim \operatorname{Kern}\varphi\,,$$

wie behauptet. ■

Wir werden diese Aussage gelegentlich als den *Homomorphiesatz* bezeichnen. Der „echte“ Homomorphiesatz ist zwar eine schärfere Aussage über den Bildraum, die auch für unendlichdimensionale Vektorräume gilt, die wir mit dem uns bislang zur Verfügung stehenden Begriffsapparat aber noch nicht formulieren können und auch nicht brauchen. Der obige Satz ist eine unmittelbare Folgerung aus dem „echten“ Homomorphiesatz.

Korollar: Eine lineare Abbildung $\varphi: V \to V$ eines endlichdimensionalen Vektorraums V auf sich selbst ist genau dann injektiv, wenn sie surjektiv ist.

Beweis: φ ist genau dann injektiv, wenn $\dim \operatorname{Bild}\varphi = \dim V$ ist, und genau dann surjektiv, wenn $\dim \operatorname{Kern}\varphi = 0$ ist. Damit folgt das Korollar sofort aus dem Satz. ■

Man beachte, daß es in diesem Korollar sehr wesentlich ist, daß wir von einem *endlichdimensionalen* Vektorraum ausgehen: Für den Vektorraum V aller reeller Polynome ist die Abbildung

$$\varphi: V \to V; \quad \sum_{i=0}^{d} a_i X^i \mapsto \sum_{i=0}^{d} a_i X^{2i}$$

linear *(warum?)* und injektiv, aber nicht surjektiv. Umgekehrt ist die Ableitung

$$\psi: V \to V; \quad f \mapsto f'$$

linear und surjektiv, aber nicht injektiv.

§2: Matrizen und lineare Gleichungssysteme

Die abstrakte Art und Weise, wie wir Vektoren und lineare Abbildungen bisher betrachtet haben, hat zwar den Vorteil, daß wir damit viele Probleme behandeltn können, die nichts mit den gewohnten anschaulichen Vektoren zu tun haben; sie hat aber auch den Nachteil, daß wir bislang noch sehr wenige nichttriviale Beispiele konkret durchrechnen können. Dieser Paragraph soll die wichtigsten Hilfsmittel zum Rechnen mit endlichdimensionalen Vektorräumen bereitstellen.

a) Abbildungsmatrizen

Basen sind nicht nur nützlich, um Vektoren darzustellen, sie können auch den Umgang mit linearen Abbildungen vereinfachen. Der Grund liegt im folgenden Lemma:

Lemma: V und W seien k-Vektorräume, und $\mathcal{B}$ sei eine Basis von V. Dann ist jede lineare Abbildung $\varphi: V \to W$ eindeutig bestimmt durch die Bilder $\varphi(\vec{b})$ der Basisvektoren $\vec{b} \in \mathcal{B}$. Umgekehrt läßt sich auch jede Abbildung $\varphi: \mathcal{B} \to W$ fortsetzen zu einer linearen Abbildung $\varphi: V \to W$

Beweis: Jeder Vektor $\vec{v} \in V$ läßt sich in eindeutiger Weise als Linearkombination

$$\vec{v} = \lambda_1\vec{b}_1 + \cdots + \lambda_n\vec{b}_n \quad \text{mit} \quad \vec{b}_1, \dots, \vec{b}_n \in \mathcal{B}$$

darstellen, und für eine lineare Abbildung φ muß dann

$$\varphi(\vec{v}) = \lambda_1\varphi(\vec{b}_1) + \cdots + \lambda_n\varphi(\vec{b}_n)$$

sein. ■

Besonders nützlich ist dies im Fall endlichdimensionaler Vektorräume.

Sei etwa V ein m-dimensionaler k-Vektorraum und W ein n-dimensionaler;

$$\mathcal{B} = \{\vec{b}_1, \dots, \vec{b}_m\} \subset V \quad \text{und} \quad \mathcal{C} = \{\vec{c}_1, \dots, \vec{c}_n\} \subset W$$

seien Basen.

Eine lineare Abbildung $\varphi: V \to W$ ist, wie wir gerade gesehen haben, eindeutig bestimmt durch die Bilder der Basisvektoren $\vec{b}_j$; diese wiederum lassen sich als Linearkombinationen der Basisvektoren $\vec{c}_i$ schreiben:

$$\varphi(\vec{b}_j) = a_{1j}\vec{c}_1 + a_{2j}\vec{c}_2 + \cdots + a_{nj}\vec{c}_n \quad \text{mit} \quad a_{ij} \in k\,.$$

Damit ist φ eindeutig bestimmt durch die $n \cdot m$ Skalare a_{ij}; wir fassen diese zusammen zu einer *Matrix:*

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

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

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

Die Matrix A zur linearen Abbildung φ bezeichen wir als *Abbildungsmatrix* von φ bezüglich der Basen $\mathcal{B}$ und $\mathcal{C}$; wenn $\mathcal{B}$ und $\mathcal{C}$ vorgegeben sind, gibt es offensichtlich für jede Matrix $A \in k^{n\times m}$ eine lineare Abbildung $\varphi: V \to W$ mit A als Abbildungsmatrix, nämlich diejenige lineare Abbildung, für die

$$\varphi(\vec{b}_j) = a_{1j}\vec{c}_1 + a_{2j}\vec{c}_2 + \cdots + a_{nj}\vec{c}_n$$

ist. Bei gegebenen Basen entsprechen sich lineare Abbildungen und Matrizen also eineindeutig (d.h. umkehrbar eindeutig: Zu jeder linearen Abbildung gibt es *genau* eine Matrix und zu jeder Matrix *genau* eine lineare Abbildung).

Als Beispiel betrachten wir den Vektorraum V aller reeller Polynome vom Grad höchstens vier und den Vektorraum W aller reeller Polynome vom Grad höchstens drei zusammen mit der linearen Abbildung

$$\varphi: V \to W; \quad f \mapsto f'\,.$$

Bevor wir eine Abbildungsmatrix berechnen können, brauchen wir zunächst Basen der beiden Vektorräume; am einfachsten nehmen wir dazu

$$\mathcal{B} = \{1, X, X^2, X^3, X^4\} \subset V \quad \text{und} \quad \mathcal{C} = \{1, X, X^2, X^3\} \subset W\,.$$

Dann ist

$$\begin{aligned} \varphi(1) &= 0 &&= 0\cdot 1 + 0\cdot X + 0\cdot X^2 + 0\cdot X^3 \\ \varphi(X) &= 1 &&= 1\cdot 1 + 0\cdot X + 0\cdot X^2 + 0\cdot X^3 \\ \varphi(X^2) &= 2X &&= 0\cdot 1 + 2\cdot X + 0\cdot X^2 + 0\cdot X^3 \\ \varphi(X^3) &= 3X^2 &&= 0\cdot 1 + 0\cdot X + 3\cdot X^2 + 0\cdot X^3 \quad \text{und} \\ \varphi(X^4) &= 4X^3 &&= 0\cdot 1 + 0\cdot X + 0\cdot X^2 + 4\cdot X^3\,, \end{aligned}$$

die Abbildungsmatrix ist also

$$A = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 0 & 4 \end{pmatrix} \in \mathbb{R}^{4\times 5}\,.$$

Hier wie auch im allgemeinen Fall stehen in den *Spalten* der Abbildungsmatrix die Koeffizienten der Bilder der Basisvektoren von V, ausgedrückt bezüglich der Basis von W.

b) Rechenregeln für Matrizen

Wir haben Matrizen eingeführt, um mit linearen Abbildungen konkret rechnen zu können; als erstes sollten wir uns dazu überlegen, wie man mit Matrizen rechnen kann.

Zu zwei $n \times m$-Matrizen

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

können wir deren Summe

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

definieren, und für einen Skalar $\lambda \in k$ auch das Produkt

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

Das legt die Vermutung nahe, daß $k^{n\times m}$ mit diesen beiden Verknüpfungen ein Vektorraum sein könnte, und in der Tat gilt:

Lemma: $k^{n\times m}$ ist ein k-Vektorraum der Dimension nm.

Beweis: Beide Rechenoperationen sind so definiert, daß, wenn wir den ij-Eintrag für sich alleine betrachten, dort die entsprechende Rechenoperation im Grundkörper k ausgeführt wird. Da für die Rechenoperationen im Grundkörper alle bei der Definition eines Vektorraums geforderten Rechenregeln gelten, gelten sie auch in $k^{n\times m}$.

Beispielsweise ist also

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

und

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

Nullelement der Addition ist natürlich die Nullmatrix

$$\begin{pmatrix} 0 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 \end{pmatrix},$$

deren sämtliche Einträge null sind, und

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

Schließlich müssen wir uns noch überlegen, daß $k^{n\times m}$ die Dimension nm hat; wir wir am Ende von §1*h)* gesehen haben, ist das gleichbedeutend damit, daß es eine Basis aus nm Matrizen gibt.

Als eine solche Basis wählen wir die Menge aller Matrizen E_{ij}, die so definiert sind, daß E_{ij} an der Stelle ij eine Eins stehen hat und sonst lauter Nullen.

In $k^{4\times 5}$ wäre also etwa $E_{23} = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$.

In einem beliebigen $k^{n\times m}$ läßt sich jede Matrix eindeutig als Linearkombination der E_{ij} schreiben, denn

$$\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1m} \\ a_{21} & a_{22} & \dots & a_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nm} \end{pmatrix} = \sum_{i=1}^{n}\sum_{j=1}^{m} a_{ij}E_{ij} \, .$$

Also ist dies eine Basis von $k^{n\times m}$, und $\dim_k k^{n\times m} = nm$. ■

Auch für zwei lineare Abbildungen $\varphi, \psi: V \to W$ können wir eine Summe definieren, und für $\lambda \in k$ auch ein Produkt $\lambda\varphi$ durch

$$\varphi + \psi: \begin{cases} V \to W \\ \vec v \mapsto \varphi(\vec v) + \psi(\vec v) \end{cases} \quad \text{und} \quad \lambda\varphi: \begin{cases} V \to W \\ \vec v \mapsto \lambda\varphi(\vec v) \end{cases} ;$$

sind V und W endlichdimensional und sind A, B die Abbildungsmatrizen von φ, ψ, so hat $\varphi + \psi$ offenbar die Abbildungsmatrix $A + B$ und λA ist die von $\lambda\varphi$.

Auch hier ist klar, daß die sämtlichen linearen Abbildungen $V \to W$ einen Vektorraum bilden, da einfach für jeden Vektor $\vec v \in V$ die Vektorraumoperationen von W auf die Bildvektoren angewendet werden; dieser Vektorraum wir mit $\mathrm{Hom}_k(V, W)$ bezeichnet nach dem Wort *Homomorphismus,* das man gelegentlich anstelle von *lineare Abbildung* gebraucht.

Im endlichdimensionalen Fall hat $\mathrm{Hom}_k(V, W)$ wegen der eineindeutigen Entsprechung von linearen Abbildungen und Matrizen als Dimension das Produkt der Dimensionen von V und von W. Die Basismatrix $E_{ij} \in k^{n\times m}$ entspricht dabei bezüglich der Basen $\mathcal{B}$ von V und $\mathcal{C}$ von W jener linearen Abbildung, die alle Vektoren aus $\mathcal{B}$ mit Ausnahme des j-ten auf den Nullvektor abbildet; der j-te Basisvektor geht auf den i-ten Basisvektor aus $\mathcal{C}$. Bei den linearen Abbildungen $k^5 \to k^4$ etwa

entspräche obige Matrix E_{23} der linearen Abbildung

$$k^5 \to k^4; \quad \begin{pmatrix} u \\ v \\ w \\ x \\ y \end{pmatrix} \mapsto \begin{pmatrix} 0 \\ w \\ 0 \\ 0 \end{pmatrix} .$$

Lineare Abbildungen lassen sich nicht nur addieren und mit Skalaren multiplizieren; sie lassen sich auch, wie alle Abbildungen, hintereinanderausführen: Sind $\psi: U \to V$ und $\varphi: V \to W$ lineare Abbildungen, so ist auch

$$\varphi \circ \psi: U \to W; \quad \vec v \mapsto \varphi\bigl(\psi(\vec v)\bigr)$$

eine lineare Abbildung.

Falls alle beteiligten Vektorräume endlichdimensional sind, können wir endliche Basen wählen; das seien etwa

$$\mathcal{B} = \{\vec b_1, \dots, \vec b_m\} \subset U, \quad \mathcal{C} = \{\vec c_1, \dots, \vec c_n\} \subset V$$

und

$$\mathcal{D} = \{\vec d_1, \dots, \vec d_p\} \subset W ,$$

d.h.

$$\dim_k U = m, \quad \dim_k V = n \quad \text{und} \quad \dim_k W = p .$$

Dann haben wir Abbildungsmatrizen $A \in k^{p\times n}$ von φ und $B \in k^{n\times m}$ von ψ; wir wollen die Abbildungsmatrix $C \in k^{p\times m}$ von $\varphi \circ \psi: U \to W$ berechnen.

Nach Definition der Abbildungsmatrizen $A = (a_{ij})$ von φ und $B = (b_{j\ell})$ von ψ ist

$$\begin{aligned} \varphi \circ \psi(\vec b_i) = \varphi\bigl(\psi(\vec b_i)\bigr) &= \varphi\left(\sum_{j=1}^{n} b_{ji}\vec c_j\right) = \sum_{j=1}^{n} b_{ji}\varphi(\vec c_j) \\ &= \sum_{j=1}^{n} b_{ji}\sum_{\ell=1}^{p} a_{\ell j}\vec d_\ell = \sum_{\ell=1}^{p}\left(\sum_{j=1}^{n} a_{\ell j}b_{ji}\right)\vec d_\ell \, . \end{aligned}$$

Für die Abbildungsmatrix $C = (c_{i\ell})$ von $\varphi \circ \psi$ ist nach Definition

$$\varphi \circ \psi(\vec{b}_i) = \sum_{\ell=1}^{p} c_{\ell i}\vec{d}_\ell \,,$$

also ist

$$c_{\ell i} = \sum_{j=1}^{n} a_{\ell j} b_{ji} \,.$$

Definition: Für zwei Matrizen $A = (a_{i\ell}) \in k^{p\times n}$ und $B = (b_{\ell j}) \in k^{n\times m}$ bezeichnen wir die Matrix $C = (c_{ij}) \in k^{p\times m}$ mit

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

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

Für praktische Rechnungen empfiehlt es sich als Eselsbrücke, den zweiten Faktor des Produkts höher zu stellen nach dem Schema

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

dadurch behält man den Überblick, welcher Rechenschritt jeweils als nächster auszuführen ist.

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,