Besonders nützlich ist dies im Fall endlichdimensionaler Vektorräume.

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

$$\mathcal{B} = (\vec{b}_1, \dots, \vec{b}_m) \quad \text{und} \quad \mathcal{C} = (\vec{c}_1, \dots, \vec{c}_n)\,.$$

Eine lineare Abbildung $\varphi\colon V \to W$ 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:

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

Somit ist φ bei gegebenen Basen 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 (geordneten) Basen $\mathcal{B}$ und $\mathcal{C}$; für eine lineare Abbildung $\varphi\colon V \to V$, deren Zielraum gleich derUrbildmenge ist, wählen wir im allgemeinen nur eine Basis $\mathcal{B}$ von V, setzen also $\mathcal{C} = \mathcal{B}$, und reden dann von der Abbildungsmatrix bezüglich der Basis $\mathcal{B}$.

Wenn $\mathcal{B}$ und $\mathcal{C}$ vorgegeben sind, gibt es offensichtlich für jede Matrix $A \in k^{n\times m}$ eine lineare Abbildung $\varphi\colon V \to W$ mit A als Abbildungsmatrix, nämlich diejenige lineare Abbildung, für die

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

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

Ein wesentlicher Punkt ist hier, daß es sich bei einer linearen Abbildung $\varphi\colon V \to W$ im allgemeinen um eine Abbildung zwischen *unendlichen* Mengen handelt und solche Abbildungen nur selten mit endlichem Aufwand beschrieben werden können. (Wie sieht etwa eine „allgemeine" Abbildung $\varphi\colon \mathbb{R} \to \mathbb{R}$ aus?) Eine lineare Abbildung zwischen endlichdimensionalen Vektorräumen ist, wie wir gerade gesehen haben, nach Wahl von Basen durch die endlich vielen Einträge der Abbildungsmatrix eindeutig bestimmt und damit einer algorithmischen Behandlung zugänglich.

Matrizen als zweidimensionale Zahlenschemata sind natürlich erheblich älter als Vektorräume und lineare Abbildungen; erste Spuren aus dem zweiten vorchristlichen Jahrhundert finden sich bereits in den *Neun Büchern der Rechenkunst* 九章算術 aus der chinesischen Han-Dynastie. Rechenregeln für den Umgang mit Matrizen tauchen ab dem 16. Jahrhundert bei den verschiedensten Autoren auf.

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

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

Bevor wir eine Abbildungsmatrix berechnen können, brauchen wir zunächst Basen der beiden Vektorräume, zum Beispiel die „üblichen" Basen

$$\mathcal{B} = (1, X, X^2, X^3, X^4) \quad \text{und} \quad \mathcal{C} = (1, X, X^2, X^3)\,.$$

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

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

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: Eigentlich gibt es nichts zu beweisen, denn wir haben einfach die Elemente aus dem Vektorraum $k^{(nm)}$ anders hingeschrieben, ohne daß dabei an den Rechenoperationen etwas zu ändern. Für diejenigen, die den Umgang mit Vektorraumaxiomen und das Rechnen mit Matrizen

noch etwas üben wollen, sei aber trotzdem noch einmal ein ausführlicher Beweis gegeben:

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

die Nullmatrix, so müssen offensichtlich alle λ_{ij} verschwinden, die E_{ij} sind also auch linear unabhängig. Damit bilden sie eine Basis von $k^{n\times m}$, und $\dim_k k^{n\times m} = nm$.

■

Die Basis mit den Matrizen E_{ij} ist zwar sicherlich die einfachste Basis für den Vektorraum aller $n\times m$-Matrizen, aber nicht immer die beste: Matrizen bieten sich beispielsweise auch an, um digitalisierte Bilder darzustellen, und zumindest in digitalen Kameras oder Scannern entsteht das Bild wirklich als eine Matrix von Grauwerten *bzw.* als drei Matrizen von Farb- oder sonstigen Werten, dargestellt in der Basis aus den E_{ij}. Für die Übertragung oder Speicherung ist das aber selten optimal, da hier für jede Komponente der Basisdarstellung dieselbe Genauigkeit erforderlich ist. Daher werden die Bilder etwa für die Speicherung im JPEG-Format zunächst in eine andere Basis umgerechnet, bezüglich derer viele Koeffizienten nahe bei Null liegen. Bei der Diskretisierung und Quantisierung enstehen dann viele Koeffizienten, für die nur wenige oder gar keine Bit benötigt werden, was im Zusammenspiel mit anderen Verfahren wie *run length encoding* und HUFFMAN-Codierung zu Komprimierungsfaktoren um die zwanzig oder dreißig ohne nennenswerten Qualitätsverlust führt.

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

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

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

$$\varphi\circ\psi\colon U \to W; \quad \vec v \mapsto \varphi\big(\psi(\vec v)\big)$$

eine lineare Abbildung.

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

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

und

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

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 $\varphi\circ\psi\colon U \to W$ berechnen.

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

$$\varphi\circ\psi(\vec b_i) = \varphi\big(\psi(\vec b_i)\big) = \varphi\left(\sum_{j=1}^n b_{ji}\vec c_j\right) = \sum_{j=1}^n b_{ji}\varphi(\vec c_j)$$
$$= \sum_{j=1}^n b_{ji}\sum_{\ell=1}^p a_{\ell j}\vec d_\ell = \sum_{\ell=1}^p\left(\sum_{j=1}^n a_{\ell j}b_{ji}\right)\vec d_\ell .$$

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

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

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

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

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

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

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

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

automatisch erfüllt.

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

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

ohne daß wir ein einziges Matrixprodukt explizit ausrechnen mußten.

Genauso folgen auch die Rechenregeln

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

aus den entsprechenden Rechenregeln

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

die sich wiederum leicht und ohne Rechnung überprüfen lassen:

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

und

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

Da in der obigen Formel für die Matrixmultiplikation alle $b_{\ell j}$ linear in den Ausdrücken für c_{ij} vorkommen *usw.*, hätten sich die beiden letzen Rechenregeln auch einfach direkt nachrechnen lassen. Ebenfalls durch direktes Nachrechnen überzeugt man sich von der Formel

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

folgt wohl am einfachsten durch direktes Nachrechnen und für die *Einheitsmatrix*

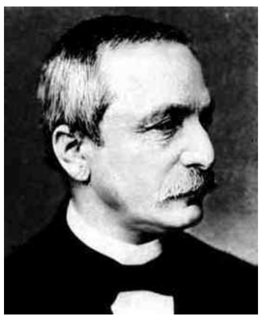
$$E = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{pmatrix} \in k^{n\times n}$$

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

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

ist. Den Eintrag an der Stelle ij der Einheitsmatrix bezeichnet man als das KRONECKER-δ:

$$\delta_{ij} = \begin{cases} 1 & \text{falls } i = j \\ 0 & \text{falls } i \neq j \end{cases} .$$



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

Bei den reellen Zahlen und auch sonst in jedem Körper gibt es zu jedem Element $a \neq 0$ ein inverses Element a^{-1}, so daß $aa^{-1} = a^{-1}a = 1$ ist. Bei Matrizen muß es das selbst für quadratische Matrizen nicht geben: Für eine beliebige Matrix $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ ist

$$\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} a & b \\ 0 & 0 \end{pmatrix}$$

und

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} a & 0 \\ c & 0 \end{pmatrix} ,$$

was beides offensichtlich nie die Einheitsmatrix sein kann.

Definition: Eine $n \times n$-Matrix $A \in k^{n\times n}$ heißt *invertierbar,* wenn es eine Matrix $B \in k^{n\times n}$ gibt, so daß $AB = BA = E$ ist. B heißt inverse Matrix von A; in Zeichen $B = A^{-1}$.

(Es wäre theoretisch möglich, Invertierbarkeit auch für nicht-quadratische Matrizen zu definieren, aber das hat keinen sonderlichen Nutzen.)

Um zu sehen, wann eine Matrix $A \in k^{n\times n}$ invertierbar ist, betrachten wir wieder die Situation bei den linearen Abbildungen: Zu einer linearen Abbildung $\varphi: k^n \to k^n$ gibt es genau dann eine Umkehrabbildung $\psi: k^n \to k^n$, so daß $\varphi \circ \psi$ und $\psi \circ \varphi$ beide die identische Abbildung sind, wenn φ bijektiv ist. Nach dem Korollar am Ende von §1*i)* ist dies genau dann der Fall, wenn φ surjektiv ist, wenn also das Bild von φ Dimension n hat. Dieses Bild wird aber erzeugt von den Bildern der Einheitsvektoren, und das sind gerade die Spalten der Abbildungsmatrix. Diese n Vektoren erzeugen genau dann ganz k^n, wenn sie linear unabhängig sind.

Definition: Der (Spalten-)Rang einer Matrix $A \in k^{n\times m}$ ist die maximale Anzahl linear unabhängiger Spaltenvektoren von A.

Nach obiger Diskussion gilt daher

Lemma: Eine Matrix $A \in k^{n\times n}$ ist genau dann invertierbar, wenn sie Rang n hat. Die inverse Matrix $B = A^{-1}$ ist sowohl durch die Bedingung $AB = E$ als auch durch die Bedingung $BA = E$ eindeutig bestimmt.

Die Eindeutigkeit der inversen Matrix folgt dabei natürlich aus der Eindeutigkeit der Umkehrabbildung. Ebenfalls klar ist das folgende

Lemma: Sind $A, B \in k^{n\times n}$ invertierbar, so auch ihr Produkt, und $(AB)^{-1} = B^{-1}A^{-1}$.

Beweis: $(AB)(B^{-1}A^{-1}) = A(BB^{-1})A^{-1} = AEA^{-1} = AA^{-1} = E$; also ist AB invertierbar und $B^{-1}A^{-1}$ ist die inverse Matrix. ■

Man beachte, daß im allgemeinen $A^{-1}B^{-1}$ *nicht* invers zu AB ist: Für

$$A = \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & 0 \\ 2 & 1 \end{pmatrix} \quad \text{und} \quad AB = \begin{pmatrix} 5 & 2 \\ 2 & 1 \end{pmatrix}$$

sind

$$A^{-1} = \begin{pmatrix} 1 & -2 \\ 0 & 1 \end{pmatrix} \quad \text{und} \quad B^{-1} = \begin{pmatrix} 1 & 0 \\ -2 & 1 \end{pmatrix}$$

die inversen Matrizen, und Multiplikation zeigt, daß

$$(AB)^{-1} = \begin{pmatrix} 1 & -2 \\ -2 & 5 \end{pmatrix} \neq \begin{pmatrix} 5 & -2 \\ -2 & 1 \end{pmatrix} = A^{-1}B^{-1}$$

ist. Insbesondere unterscheidet sich auch

$$ABA^{-1}B^{-1} = \begin{pmatrix} 21 & -8 \\ 8 & -3 \end{pmatrix}$$

deutlich von der Einheitsmatrix.

c) Matrixdarstellung der komplexen Zahlen

Die komplexen Zahlen bilden mit ihrer üblichen Addition und der Einschränkung der üblichen Multiplikation zu einer Abbildung

$$\cdot : \begin{cases} \mathbb{R} \times \mathbb{C} \to \mathbb{C} \\ (r, z) \mapsto rz \end{cases}$$

einen $\mathbb{R}$-Vektorraum und die Abbildung

$$\mathbb{C} \to \mathbb{C}; \quad z \mapsto cz$$

ist für jede komplexe Zahl $c = a + ib \in \mathbb{C}$ insbesondere eine lineare Abbildung von $\mathbb{R}$-Vektorräumen. Wählen wir $\{1, i\}$ als $\mathbb{R}$-Basis von $\mathbb{C}$, so bildet sie die beiden Basisvektoren 1 und i ab auf

$$c \cdot 1 = a + bi \quad \text{und} \quad c \cdot i = ai + bi^2 = -b + ai\,;$$

sie hat also die Abbildungsmatrix

$$\begin{pmatrix} a & -b \\ b & a \end{pmatrix}.$$

Die Hintereinanderausführung zweier solcher Abbildungen entspricht der Multiplikation der entsprechenden komplexen Zahlen; insbesondere

gehört also das Produkt $(a + ib)(a' + ib')$ zweier komplexer Zahlen zur Produktmatrix

$$\begin{pmatrix} a & -b \\ b & a \end{pmatrix}\begin{pmatrix} a' & -b' \\ b' & a' \end{pmatrix}.$$

Der Körper der komplexen Zahlen kann damit auch identifiziert werden mit der Menge aller reeller 2×2-Matrizen der obigen Form mit der Matrixaddition und dem Matrixprodukt. Man beachte, daß das Produkt zweier Matrizen dieser speziellen Form kommutativ ist, denn die Multiplikation komplexer Zahlen ist kommutativ.

Betrachten wir speziell den Fall, daß c den Betrag eins hat. Dann ist $|cz| = |c| \cdot |z| = |z|$, und allgemeiner ist für zwei beliebige komplexe Zahlen z, w auch

$$|cz - cw| = |c(z - w)| = |c| \cdot |z - w| = |z - w|\,,$$

der EUKLIDische Abstand zwischen den Bildpunkten cz und zw ist also gleich dem Abstand zwischen z und w. Die Abbildung $z \mapsto cz$ ist somit eine Kongruenzabbildung. Da sie für $c \neq 1$ den Nullpunkt als einzign Fixpunkt hat, ist sie entweder eine Drehung oder eine Drehspiegelung. Letzteres kommt nicht in Frage, da man kontinuierlich auf dem Einheitskreis von eins zu c gehen kann, also haben wir eine Drehung um einen Winkel φ.

Zur Bestimmung von φ können wir ausnutzen, daß diese Drehung den Punkt Eins in den Punkt c überführt: Betrachten wir dazu das rechtwinklige Dreieck mit Ecken $0, 1$ und $c = a + ib$! Seine Hypothenuse hat die Länge eins, seine Katheten sind a und b. Nach Definition der Winkelfunktionen als Ankathete *bzw.* Gegenkathete durch Hypothenuse ist offensichtlich $a = \cos\varphi$ und $b = \sin\varphi$. Somit ist $c = \cos\varphi + i\sin\varphi$.

Schreiben wir für einen beliebigen Winkel φ

$$c_\varphi = a_\varphi + ib_\varphi \quad \text{mit} \quad a_\varphi = \cos\varphi \quad \text{und} \quad b_\varphi = \sin\varphi\,,$$

so ist offensichtlich $c_\varphi c_\psi = c_{\varphi+\psi}$, denn die Hintereinanderausführung zweier Drehungen um den Nullpunkt ist eine Drehung mit der Summe der beiden Drehwinkel. Ausmultipliziert ergibt dies

$$c_{\varphi+\psi} = (a_\varphi + ib_\varphi)(a_\psi + ib_\psi) = (a_\varphi a_\psi - b_\varphi b_\psi) + i(a_\varphi b_\psi + a_\psi b_\varphi)\,,$$

d.h.
$$\cos(\varphi+\psi) = \cos\varphi\cos\psi - \sin\varphi\sin\psi \quad \text{und}$$
$$\sin(\varphi+\psi) = \sin\varphi\cos\psi + \cos\varphi\sin\psi\,;$$
wir haben also die Additionstheoreme für Sinus und Kosinus gefunden.

Außerdem zeigt die Beziehung $c_\varphi c_\psi = c_{\varphi+\psi}$, daß sich c_φ als Funktion von φ wie eine Potenz verhält; wir schreiben deshalb, im Augenblick noch formal,
$$c_\varphi = e^{i\varphi} \underset{\text{def}}{=} \cos\varphi + i\sin\varphi\,.$$

Offensichtlich läßt sich jede komplexe Zahl in der Form $z = re^{i\varphi}$ schreiben, wobei $r = |z|$ der Betrag von z ist. Für $z = 0$ können wir für φ jeden beliebigen Wert nehmen; für $z \neq 0$ ist φ modulo 2π eindeutig bestimmt. In diesem Fall bezeichnen wir $\varphi = \arg z$ als das *Argument* von z und (r, φ) als die *Polarkoordinaten* von z.

In Polarkoordinaten wird die Multiplikation komplexer Zahlen deutlich einfacher als in den bislang benutzten kartesischen Koordinaten:
$$(re^{i\varphi}) \cdot (se^{i\psi}) = rs \cdot e^{i(\varphi+\psi)}\,.$$
Dies läßt sich auch benutzen, um Wurzeln zu ziehen: Offensichtlich ist $\sqrt[n]{r} \cdot e^{i\varphi/n}$ eine n-te Wurzel aus $re^{i\varphi}$; weitere Wurzeln sind die Zahlen $\sqrt[n]{r} \cdot e^{i(\varphi+2k\pi)/n}$ für $k = 1, \ldots, n-1$. (Für $k = n$ bekommen wir wieder dieselbe Zahl wie für $k = 0$.)

Beispielsweise ist in Polarkoordinaten $i = e^{\pi i/2}$, da i aus der Eins durch Drehung um 90° oder $\frac{\pi}{2}$ entsteht. Also ist $e^{\pi i/4}$ eine Quadratwurzel aus i. Der Winkel $\frac{\pi}{4}$ oder 45° tritt in gleichschenkligen rechtwinkligen Dreiecken auf, die man auch auffassen kann als entlang der Diagonale halbierte Quadrate; da die Diagonale eines Quadrats das $\sqrt{2}$-fache der Seite ist, sind also Sinus und Cosinus gleich $\frac{1}{\sqrt{2}} = \frac{\sqrt{2}}{2}$ und somit ist
$$\cos\frac{\pi}{4} + i\sin\frac{\pi}{4} = \frac{\sqrt{2}}{2} + i\frac{\sqrt{2}}{2}$$
eine Quadratwurzel aus i ist. Die andere ist natürlich einfach das Negative davon.

Nicht nur komplexe Zahlen lassen sich mit Matrizen identifizieren; Ganz entsprechend kann man auch die Elemente der Körper $\mathbb{F}_{2^n}$ mit $n \times n$-Matrizen über $\mathbb{F}_2$ identifizieren. Nimmt man etwa $1, \alpha$ als Basisvektoren des $\mathbb{F}_2$-Vektorraums $\mathbb{F}_4$, so entsprechen die vier Elemente $0, 1, \alpha$ und $\alpha + 1$ des Körpers $\mathbb{F}_4$ den Matrizen
$$\begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}, \quad \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix} \quad \text{und} \quad \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix},$$
und für $\mathbb{F}_{32}$ mit $\mathbb{F}_2$-Basis $1, \alpha, \alpha^2, \alpha^3, \alpha^4$ und Relation $\alpha^5 = \alpha^2 + 1$ entspricht α der Matrix
$$\begin{pmatrix} 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{pmatrix}.$$

d) Das Gaußsche Eliminationsverfahren

Die meisten kennen wohl aus der Schule zumindest Verfahren zur Lösung linearer Gleichungssysteme in bis zu drei Unbekannten, teilweise vielleicht auch für Systeme aus beliebig vielen Gleichungen in beliebig vielen Unbekannten.

Der GAUSS-Algorithmus, mit dem wir uns hier beschäftigen wollen, bestimmt die Lösungsmenge eines beliebigen linearen Gleichungssystems, und falls das Gleichungssystem nicht gerade eine sehr spezielle Gestalt hat, liefert er sie im allgemeinen auf die effizienteste Art und Weise.

Seine Grundidee ist sehr einfach: Im Falle einer einzigen Gleichung mit einer einzigen Unbekannten x können wir das „Gleichungssystem“
$$ax = b$$
sofort lösen: Für $a \neq 0$ ist $x = -b/a$, d.h. $\mathcal{L} = \{-b/a\}$; ansonsten gibt es für $b \neq 0$ keine Lösung, d.h. $\mathcal{L} = \emptyset$, und für $a = b = 0$ ist jedes x aus k eine Lösung, d.h. $\mathcal{L} = k$.

Das GAUSSsche Eliminationsverfahren führt ein allgemeines lineares Gleichungssystem sukzessive zurück auf solche lineare Gleichungen in einer Unbekannten, ausgehend von zwei trivialen Beobachtungen:

1. Die Lösungsmenge eines linearen Gleichungssystems ändert sich nicht, falls wir zwei Gleichungen miteinander vertauschen.
2. Die Lösungsmenge ändert sich auch dann nicht, wenn wir ein Vielfaches einer Gleichung zu einer anderen addieren, d.h. wenn wir die Gleichung

   $$\ell_j(x_1,\dots,x_m) \underset{\text{def}}{=} a_{j1}x_1 + \ldots + a_{jm}x_m = b_j$$

   ersetzen durch

   $$\ell_j(x_1,\dots,x_m) + \lambda\ell_i(x_1,\dots,x_m) = b_j + \lambda b_i\,, \qquad (*)$$

   denn unter der Nebenbedingung

   $$\ell_i(x_1,\dots,x_m) = b_i$$

   ist $(*)$ äquivalent zu $\ell_j(x_1,\dots,x_m) = 0$.

Mit Hilfe dieser beiden Beobachtungen läßt sich nun die Variablenanzahl wie folgt sukzessive reduzieren: Beginnen wir mit der Elimination von x_1. Falls x_1 im Gleichungssystem überhaupt nicht vorkommt, falls also alle $a_{i1} = 0$ sind, gibt es nichts zu tun: Wir haben ein Gleichungssystem in $x_2,\dots,x_m$, und für jede Lösung $(x_2,\dots,x_m)$ dieses Systems sowie jedes beliebige $x_1 \in k$ ist $(x_1,x_2,\dots,x_m)$ eine Lösung des ursprünglichen Systems.

Ansonsten können wir, indem wir nötigenfalls zwei Gleichungen miteinander vertauschen, annehmen, daß $a_{11} \neq 0$ ist. Dann lassen wir die erste Gleichung so stehen, wie sie ist, und ersetzen jede weitere Gleichung $\ell_j(x_1,\dots x_m) = b_j$ durch

$$\ell_j(x_1,\dots x_m) - \frac{a_{j1}}{a_{11}}\ell_1(x_1,\dots x_m) = b_j - \frac{a_{j1}}{a_{11}}b_1\,;$$

in diesen Gleichungen kommt x_1 offenbar nicht mehr vor. Wir haben somit ein Gleichungssystem in einer Variablen weniger, plus der Gleichung

$$a_{11}x_1 + a_{12}x_2 + \dots + a_{1m}x_m = b_1\,.$$

Sobald wir das Gleichungssytem für $x_2,\dots,x_m$ gelöst haben, wird diese Gleichung nach Einsetzen einer Lösung $(x_2,\dots,x_m)$ zu einer linearen Gleichung für x_1, die wir lösen können:

$$x_1 = \frac{b_1 - (a_{12}x_2 + \dots + a_{1m}x_m)}{a_{11}}\,.$$

Das Gleichungssystem für $x_2,\dots,x_m$ wird nun, falls $m > 2$ ist, nach genau derselben Methode weiterreduziert: Falls x_2 in keiner der Gleichungen vorkommt, haben wir tatsächlich ein Gleichungssystem in $x_3,\dots,x_m$, andernfalls können wir durch Vertauschen zweier Gleichungen annehmen, daß x_2 in der ersten Gleichung mit einem von Null verschiedenen Vorfaktor auftritt, und wir können wie oben x_2 aus allen weiteren Gleichungen eliminieren *usw.*

Da in jedem Eliminationsschritt ein Nenner auftritt, können die Nenner vor allem bei großem m gelegentlich schnell unübersichtlich groß werden. Obwohl es *im Prinzip* nicht notwendig ist, kann man um dies zu vermeiden noch als dritte Operation die Multiplikation einer Gleichung mit einem Körperelement (z.B. einem Hauptnenner der Koeffizienten) zulassen. Dies ist gleichbedeutend damit, daß man anstelle der Operation $(*)$ allgemeiner die Ersetzung von $\ell_j(x_1,\dots x_m)$ durch

$$\mu\ell_j(x_1,\dots x_m) + \lambda\ell_i(x_1,\dots x_m)$$

mit beliebigem $\mu \neq 0$ aus k zuläßt. ($\mu = 0$ muß hier natürlich unbedingt ausgeschlossen werden, denn sonst läßt sich die Gleichung $\ell_j(x_1,\dots,x_m) = 0$ aus dem neuen Gleichungssystem nicht mehr herleiten, d.h. wir erhalten auch „Lösungen“, die diese Gleichung nicht erfüllen.) Ein spezieller Fall der Multiplikation einer Gleichung mit einem Körperelement ist das Kürzen durch einen gemeinsamen Teiler der Koeffizienten, wodurch ein Gleichungssystem (egal ob das ursprüngliche oder ein Zwischenresultat) und vor allem der weitere Rechengang oft erheblich übersichtlicher wird.

e) Erste Beispiele

Betrachten wir dazu einige Beispiele; der Grundkörper k sei dabei jeweils der Körper der reellen Zahlen oder einer seiner Teilkörper, etwa $k = \mathbb{Q}$.

Sei zunächst

$$\begin{array}{rcrcrcr} 3x_1 & + & 2x_2 & + & x_3 & = & 5 \\ x_1 & + & 2x_2 & + & 4x_3 & = & 3 \\ 5x_1 & - & 3x_2 & + & 7x_2 & = & 19 \end{array}$$

das zu lösende Gleichungssystem. Da x_1 in der ersten Gleichung tatsächlich vorkommt, müssen wir nichts vertauschen; allerdings müssen wir ein Drittel der ersten Gleichung von der zweiten und fünf Drittel der ersten Gleichung von der dritten subtrahieren, um x_2 aus diesen beiden Gleichungen zu eliminieren, was auf das etwas unangenehme Gleichungssystem

$$\begin{array}{rcrcrcr} 3x_1 & + & 2x_2 & + & x_3 & = & 5 \\ & & \frac{4}{3}x_2 & + & \frac{11}{3}x_3 & = & \frac{4}{3} \\ & & -\frac{19}{3}x_2 & + & \frac{16}{3}x_3 & = & \frac{32}{3} \end{array}$$

führt. Solche Gleichungssysteme sind zwar nicht immer vermeidbar, aber hier hätten wir es auch einfacher haben können: Wenn wir im ursprünglichen Gleichungssystem die ersten beiden Gleichungen vertauschen, wird es zu

$$\begin{array}{rcrcrcr} x_1 & + & 2x_2 & + & 4x_3 & = & 3 \\ 3x_1 & + & 2x_2 & + & x_3 & = & 5 \\ 5x_1 & - & 3x_2 & + & 7x_2 & = & 19\,, \end{array}$$

und hier müssen wir stattdessen das Dreifache *bzw.* Fünffache der ersten Gleichung von der zweiten *bzw.* dritten subtrahieren, was auf das deutlich angenehmere Gleichungssystem

$$\begin{array}{rcrcrcr} x_1 & + & 2x_2 & + & 4x_3 & = & 3 \\ & & -4x_2 & - & 11x_3 & = & -4 \\ & & -13x_2 & - & 13x_3 & = & 4 \end{array}$$

führt. Etwas ähnliches hätten wir auch bekommen, wenn wir die letzten beiden Gleichungen des anderen Systems einfach mit drei multipliziert hätten, aber grundsätzlich ist es meistens günstiger, die Gleichung mit dem einfachsten führenden Koeffizienten an erster Stelle zu haben. Auch hier wird das Gleichungssystem zumindest optisch etwas angenehmer, wenn wir die zweite und die dritte Gleichung mit (-1) multiplizieren:

$$\begin{array}{rcrcrcr} x_1 & + & 2x_2 & + & 4x_3 & = & 3 \\ & & 4x_2 & & 11x_3 & = & 4 \\ & & 13x_2 & + & 13x_3 & = & -4 \end{array}$$

Uns interessieren zunächst nur die letzten beiden Zeilen. Diese bilden ein lineares Gleichungssystem in x_2 und x_3, aus dem wir x_2 in einer der beiden Gleichungen eliminieren möchten.

Da x_2 in der zweiten Gleichung wirklich vorkommt, subtrahieren wir $13/4$ mal diese Gleichung von der dritten und erhalten als neues System

$$\begin{array}{rcrcrcr} 3x_1 & + & 2x_2 & + & x_3 & = & 5 \\ & & 4x_2 & + & 11x_3 & = & 4 \\ & & & & -\frac{91}{4}x_3 & = & -17\,. \end{array}$$

Hier ist die letzte Gleichung eine gewöhnliche lineare Gleichung für x_3, aus der wir sofort ablesen, daß

$$x_3 = \frac{68}{91}$$

ist. Dies setzen wir in die vorletzte Gleichung ein:

$$4x_2 + \frac{748}{91} = 4$$

ist eine lineare Gleichung für x_2 mit Lösung

$$x_2 = -\frac{96}{91}\,.$$

Dies sowie x_3 setzen wir schließlich in die erste Gleichung ein:

$$x_1 + \frac{80}{91} = 3$$

hat die Lösung

$$x_1 = \frac{193}{91}\,,$$

so daß ingesamt

$$\left(\frac{193}{91}, -\frac{96}{91}, \frac{68}{91}\right)$$

die (einzige) Lösung des Gleichungssystems ist.

Als nächstes wollen wir ein Beispiel betrachten, bei dem nicht mit einer eindeutigen Lösung zu rechnen ist: Wir nehmen ein System aus drei Gleichungen in sieben Unbekannten. Zumindest intuitiv schein klar, daß man mit nur drei Bedingungen sieben Variable nicht eindeutig festlegen kann; wir wollen sehen, was der GAUSS-Algorithmus in so einer Situation liefert.

Das Gleichungssystem ist

$$\begin{array}{rcrcrcrcrcrcrcl} & & & & 5x_3 & & & & & - & 2x_6 & & & = & 3 \\ & & 4x_2 & & & + & 2x_4 & & & - & 3x_6 & - & 7x_7 & = & -2 \\ x_1 & & & & & - & 3x_4 & + & x_5 & & & & & = & 0 \end{array}$$

Hier kommt x_1 in der ersten Gleichung nicht vor, wohl aber in der dritten. Wir vertauschen daher die erste Gleichung mit der dritten und erhalten

$$\begin{array}{rcrcrcrcrcrcrcl} x_1 & & & & & - & 3x_4 & + & x_5 & & & & & = & 0 \\ & & 4x_2 & & & + & 2x_4 & & & - & 3x_6 & - & 7x_7 & = & -2 \\ & & & & 5x_3 & & & & & - & 2x_6 & & & = & 3\,. \end{array}$$

In diesem Gleichungssystem kommt x_1 nur in der ersten Gleichung vor, x_2 nicht mehr hinter der zweiten und x_3 (natürlich) nicht mehr hinter der dritten, wir haben also bereits die Treppengestalt erreicht.

Die dritte Gleichung $5x_3 - 2x_6 = 3$ hat offensichtlich keine eindeutig bestimmte Lösung: Wir können entweder nach x_3 oder nach x_6 auflösen, so daß eine der beiden Unbekannten beliebig bewählt werden kann. Wählen wir etwa für x_6 irgendein Element $\alpha \in k$, so wird

$$x_6 = \alpha, \quad x_3 = \frac{2}{5}\alpha + \frac{3}{5}\,.$$

Die zweite Gleichung wird nach Einsetzen von $x_6 = \alpha$ zu

$$4x_2 + 2x_4 - 3\alpha - 7x_7 = -2\,,$$

also zu einer Beziehung zwischen x_2, x_4 und x_7, denn α ist ein bereits fest gewähltes und somit bekanntes Element des Grundkörpers k. Wieder können wir eine der beiden Variablen willkürlich festlegen, etwa x_7 auf ein $\beta \in k$ setzen und x_4 auf ein $\gamma \in k$; wir erhalten

$$x_7 = \beta, \quad x_4 = \gamma, \quad x_2 = \frac{3}{4}\alpha + \frac{7}{4}\beta - \frac{\gamma}{2} - \frac{1}{2}\,.$$

In der ersten Gleichung schließlich müssen wir alle bereits berechneten oder festgelegten Unbekannten einsetzen und erhalten dann die Gleichung

$$x_1 - 3\gamma + x_5 = 0\,.$$

Hier können wir noch $x_5 = \delta$ beliebig festlegen und erhalten dann für die letzte noch verbleibende Variable

$$x_1 = 3\gamma - \delta\,.$$

Insgesamt hängt die Lösung dieses Gleichungssystems also ab von vier willkürlich wählbaren Körperelementen $\alpha, \beta, \gamma, \delta \in k$ oder, wie wir auch sagen werden, von vier *Parametern*. Die *Lösungsmenge* $\mathcal{L}$ ist daher die Menge

$$\left\{ \left(3\gamma - \delta, \frac{3}{4}\alpha + \frac{7}{4}\beta - \frac{\gamma}{2} - \frac{1}{2}, \frac{2}{5}\alpha + \frac{3}{5}, \gamma, \delta, \alpha, \beta\right) \,\middle|\, \alpha, \beta, \gamma, \delta \in k \right\}.$$

Im ersten Beispiel brauchten wir von den beiden Operationen des GAUSS-Algorithmus nur den Eliminationsschritt, im zweiten nur den Vertauschungsschritt. Als nächstes betrachten wir ein Beispiel, in dem beide notwendig sind, das lineare Gleichungssystem

$$\begin{array}{rcrcrcrcrcrcl} & & x_2 & & & + & 5x_4 & - & 2x_5 & + & x_6 & = & 10 \\ 2x_1 & + & 2x_2 & & & - & 4x_4 & - & 6x_5 & + & 2x_6 & = & 12 \\ 6x_1 & - & 3x_2 & + & x_3 & + & 5x_4 & & & - & 3x_6 & = & 15 \\ 4x_1 & + & x_2 & & & - & 3x_4 & & & + & x_6 & = & 0 \\ 6x_1 & - & 6x_2 & + & x_3 & + & 10x_4 & + & 12x_5 & - & 6x_6 & = & -8 \end{array}$$

Hier kommt x_1 ausgerechnet in der ersten Gleichung nicht vor, dafür aber in allen folgenden; wir müssen also die erste Gleichung mit einer der anderen vertauschen, etwa der zweiten:

$$\begin{array}{rcrcrcrcrcrcl} 2x_1 & + & 2x_2 & & & - & 4x_4 & - & 6x_5 & + & 2x_6 & = & 12 \\ & & x_2 & & & + & 5x_4 & - & 2x_5 & + & x_6 & = & 10 \\ 6x_1 & - & 3x_2 & + & x_3 & + & 5x_4 & & & - & 3x_6 & = & 15 \\ 4x_1 & - & 7x_2 & & & - & 3x_4 & & & + & x_6 & = & 0 \\ 2x_1 & - & 7x_2 & + & x_3 & + & 13x_4 & + & 12x_5 & - & 6x_6 & = & -8 \end{array}$$

Nun kommt x_1 in der zweiten Gleichung nicht mehr vor; aus der dritten, vierten *bzw.* fünften Gleichung kann x_1 eliminiert werden, indem man dreimal, zweimal *bzw.* einmal die erste Gleichung subtrahiert:

$$\begin{array}{rcrcrcrcrcrcl} 2x_1 & + & 2x_2 & & & - & 4x_4 & - & 6x_5 & + & 2x_6 & = & 12 \\ & & x_2 & & & + & 5x_4 & - & 2x_5 & + & x_6 & = & 10 \\ & & -9x_2 & + & x_3 & + & 17x_4 & + & 18x_5 & - & 9x_6 & = & -21 \\ & & -3x_2 & & & + & 5x_4 & + & 12x_5 & - & 3x_6 & = & -24 \\ & & -12x_2 & + & x_3 & + & 22x_4 & + & 30x_5 & - & 12x_6 & = & -44 \end{array}$$

Als nächstes muß x_2 aus der dritten bis fünften Gleichung eliminiert werden; da x_2 in der zweiten Gleichung mit Koeffizient eins vorkommt, müssen wir einfach jenes Vielfache der zweiten Gleichung subtrahieren, das dem jeweiligen x_2-Koeffizienten entspricht. Da hier alle diese Koeffizienten negativ sind, bedeutet dies, daß wir dasjenige Vielfache der zweiten Gleichung *addieren*, das dem Betrag des Koeffizienten entspricht:

$$\begin{array}{rcrcrcrcrcrcrcr}
2x_1 & + & 2x_2 & & & - & 4x_4 & - & 6x_5 & + & 2x_6 & = & 12 \\
 & & x_2 & & & + & 5x_4 & - & 2x_5 & + & x_6 & = & 10 \\
 & & & & x_3 & + & 62x_4 & & & & & = & 69 \\
 & & & & & & 20x_4 & + & 6x_5 & & & = & 6 \\
 & & & & x_3 & + & 82x_4 & + & 6x_5 & & & = & 76\,.
\end{array}$$

Hier muß x_3 aus der letzten Gleichung eliminiert werden; dazu muß offenbar einfach die dritte Gleichung subtrahiert werden:

$$\begin{array}{rcrcrcrcrcrcrcr}
2x_1 & + & 2x_2 & & & - & 4x_4 & - & 6x_5 & + & 2x_6 & = & 12 \\
 & & x_2 & & & + & 5x_4 & - & 2x_5 & + & x_6 & = & 10 \\
 & & & & x_3 & + & 62x_4 & & & & & = & 69 \\
 & & & & & & 20x_4 & + & 6x_5 & & & = & 6 \\
 & & & & & & 20x_4 & + & 6x_5 & & & = & 7\,.
\end{array}$$

Eigentlich sollte man hier schon sehen, was los ist, aber wir rechnen zur Veranschaulichung des GAUSS-Algorithmus trotzdem stur weiter nach Schema F: Danach muß x_4 aus der letzten Gleichung eliminiert werden durch Subtraktion der vorletzten:

$$\begin{array}{rcrcrcrcrcrcrcr}
2x_1 & + & 2x_2 & & & - & 4x_4 & - & 6x_5 & + & 2x_6 & = & 12 \\
 & & x_2 & & & + & 5x_4 & - & 2x_5 & + & x_6 & = & 10 \\
 & & & & x_3 & + & 62x_4 & & & & & = & 69 \\
 & & & & & & 20x_4 & + & 6x_5 & & & = & 6 \\
 & & & & & & & & & & 0 & = & 1\,.
\end{array}$$

Damit wird endgültig klar, daß jede Lösung $(x_1, \ldots, x_6)$ des gegebenen Gleichungssystems insbesondere auch die Gleichung $0 = 1$ erfüllen muß, d.h. die Lösungsmenge ist leer.

Nachdem wir soviel Arbeit in dieses Beispiel investiert haben, sollten wir zumindest einen Teil der Rechnungen recyclen zu einem Beispiel, in dem es Lösungen gibt. Dazu muß nur die letzte der fünf ursprünglichen Gleichungen auf der rechten Seite leicht abgeändert werden: Wir

betrachten nun das System

$$\begin{array}{rcrcrcrcrcrcrcr}
 & & x_2 & & & + & 5x_4 & - & 2x_5 & + & x_6 & = & 10 \\
2x_1 & + & 2x_2 & & & - & 4x_4 & - & 6x_5 & + & 2x_6 & = & 12 \\
6x_1 & - & 3x_2 & + & x_3 & + & 5x_4 & & & - & 3x_6 & = & 15 \\
4x_1 & + & x_2 & & & - & 3x_4 & & & + & x_6 & = & 0 \\
6x_1 & - & 6x_2 & + & x_3 & + & 10x_4 & + & 12x_5 & - & x_6 & = & \mathbf{-9}\,.
\end{array}$$

Hierauf lassen sich genau dieselben Umformungen anwenden wie oben, anstelle des Systems mit fünfter Gleichung $0 = 1$ führen diese nun aber auf

$$\begin{array}{rcrcrcrcrcrcrcr}
2x_1 & + & 2x_2 & & & - & 4x_4 & - & 6x_5 & + & 2x_6 & = & 12 \\
 & & x_2 & & & + & 5x_4 & - & 2x_5 & + & x_6 & = & 10 \\
 & & & & x_3 & + & 62x_4 & & & & & = & 69 \\
 & & & & & & 20x_4 & + & 6x_5 & & & = & 6 \\
 & & & & & & & & & & 0 & = & 0\,.
\end{array}$$

Diese letzte Gleichung ist natürlich für jedes Tupel $(x_1, \ldots, x_6)$ erfüllt. Die vorletzte gibt eine Beziehung zwischen x_4 und x_5, wir können also

$$x_5 = \alpha \in k$$

beliebig wählen und erhalten dann

$$x_4 = -\frac{3}{10}\alpha + \frac{3}{10}\,.$$

Wenn wir dies in die dritte Gleichung einsetzen, bleibt dort nur noch x_3 als Variable stehen, und aus

$$x_3 - \frac{93}{5}\alpha + \frac{93}{5} = 69$$

lesen wir sofort ab, daß

$$x_3 = \frac{93}{5}\alpha + \frac{252}{5}$$

ist. Damit gehen wir in die zweite Gleichung:

$$x_2 - \frac{7}{5}\alpha + x_6 + \frac{3}{2} = 10\,.$$

Hier können wir wieder eine der beiden noch verbliebenen Variablen auf einen beliebigen Wert setzen, etwa

$$x_6 = \beta \in k\,.$$

Dann wird

$$x_2 = \frac{7}{2}\alpha - \beta + \frac{17}{2}\,,$$

was wir schließlich zusammen mit all den anderen bereits berechneten x_i in die erste Gleichung einsetzen können:

$$2x_1 + \frac{11}{5}\alpha + \frac{79}{5} = 12$$

hat die Lösung

$$x_1 = -\frac{11}{10}\alpha - \frac{19}{10}\,.$$

Damit hängt also die allgemeine Lösung dieses linearen Gleichungssystems von zwei frei wählbaren Parametern α und β ab. Sie wird geringfügig übersichtlicher, wenn wir α als $\alpha = 10\gamma$ schreiben; dann ist $\mathcal{L}$ gleich der Menge

$$\left\{\left(-11\gamma - \tfrac{19}{10}, 35\gamma - \beta + \tfrac{17}{2}, 186\gamma + \tfrac{252}{5}, -3\gamma + \tfrac{3}{10}, 10\gamma, \beta\right) \;\middle|\; \beta, \gamma \in k\right\}.$$

Als nächstes Beispiel wollen wir ein System betrachten, daß von zwar festen, aber nicht numerisch gegebenen Parametern abhängt: Wir betrachten das Gleichstromnetzwerk aus Abbildung zwölf mit bekannten Widerständen $R_1, \ldots, R_5$ und bekanntem Eingangs- und Ausgangsstrom I; gesucht sind die Ströme $I_1, \ldots, I_5$.

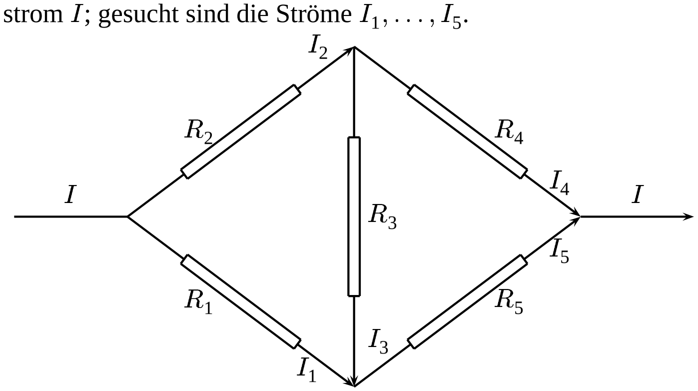


Abb. 12: Ein Gleichstromnetzwerk

Nach den KIRCHHOFFschen Gesetzen müssen sich zunächst an allen Knoten die eingehenden und die ausgehenden Ströme zu Null ergänzen, d.h. wir erhalten die Gleichungen

$$I_1 + I_2 = I = I_4 + I_5$$

für Anfang und Ende sowie

$$I_2 = I_3 + I_4 \quad \text{und} \quad I_1 + I_3 = I_5$$

für oben und unten. Außerdem müssen sich die Spannungen in den beiden Dreiecken zu Null summieren; nach dem OHMschen Gesetz führt dies auf die beiden Gleichungen

$$R_2I_2 + R_3I_3 - R_1I_1 = 0 \quad \text{und} \quad R_3I_3 + R_5I_5 - R_4I_4 = 0\,.$$

Ordnen wir dies nach den Strömen I_ν, erhalten wir also das lineare Gleichungssystem

$$\begin{array}{rcrcrcrcrcl}
I_1 & + & I_2 & & & & & & & = & I \\
 & & & & & & I_4 & + & I_5 & = & I \\
 & & I_2 & - & I_3 & - & I_4 & & & = & 0 \\
I_1 & & & + & I_3 & & & - & I_5 & = & 0 \\
-R_1I_1 & + & R_2I_2 & + & R_3I_3 & & & & & = & 0 \\
 & & & & -R_3I_3 & + & R_4I_4 & - & R_5I_5 & = & 0\,.
\end{array}$$

Die Elimination von I_1 aus allen Gleichungen ab der zweiten ist einfach: Wir müssen die erste Gleichung von der vierten subtrahieren und ihr R_1-faches zur fünften addieren; das neue System ist

$$\begin{array}{rcrcrcrcrcr}
I_1 & + & I_2 & & & & & & & = & I \\
 & & & & & & I_4 & + & I_5 & = & I \\
 & & I_2 & - & I_3 & - & I_4 & & & = & 0 \\
 & & -I_2 & + & I_3 & & & - & I_5 & = & -I \\
 & & (R_1 + R_2)I_2 & + & R_3I_3 & & & & & = & R_1I \\
 & & & & -R_3I_3 & + & R_4I_4 & - & R_5I_5 & = & 0\,,
\end{array}$$

Da I_2 in der zweiten Gleichung nicht vorkommt, vertauschen wir diese mit der dritten; danach können wir letztere zur vierten addieren und ihr $(R_1 + R_2)$-faches von der fünften subtrahieren, um I_2 aus allen weiteren Gleichungen zu eliminieren.

Damit weiterhin jede Gleichung in eine Zeile paßt, setzen wir zur Abkürzung

$$R_{12} \underset{\text{def}}{=} R_1 + R_2 \quad \text{und} \quad R_{123} \underset{\text{def}}{=} R_1 + R_2 + R_3$$

und erhalten

$$\begin{array}{rcrcrcrcrcr}
I_1 & + & I_2 & & & & & & & = & I \\
 & & I_2 & - & I_3 & - & I_4 & & & = & 0 \\
 & & & & & & I_4 & + & I_5 & = & I \\
 & & & & & & -I_4 & + & -I_5 & = & -I \\
 & & & & R_{123}I_3 & + & R_{12}I_4 & & & = & R_1 I \\
 & & & & -R_3I_3 & + & R_4I_4 & - & R_5I_5 & = & 0\,.
\end{array}$$

Zur Elimination von I_3 bietet sich an, die dritte Gleichung, in der I_3 nicht vorkommt, mit der sechsten zu vertauschen (diese hat einen einfacheren I_3-Koeffizienten als die fünfte), und dann das R_{123}/R_3-fache dieser Gleichung zur fünften zu addieren. Dazu müssen wir uns allerdings zunächst überlegen, ob das überhaupt zulässig ist: Falls $R_3 = 0$ ist, dürfen wir natürlich nicht durch R_3 dividieren; wenn dieser Fall nicht ausgeschlossen werden kann, muß er also ab hier getrennt behandelt werden.

Im vorliegenden Beispiel wollen wir aber davon ausgehen, daß alle fünf Widerstände tatsächlich vorhanden sind, so daß alle R_i positive reelle Zahlen sind. Dann können wir durch R_3 dividieren und das System wird zu

$$\begin{array}{rcrcrcrcrcr}
I_1 & + & I_2 & & & & & & & = & I \\
 & & I_2 & - & I_3 & - & I_4 & & & = & 0 \\
 & & & & -R_3I_3 & + & R_4I_4 & - & R_5I_5 & = & 0 \\
 & & & & & & -I_4 & + & -I_5 & = & -I \\
 & & & & & & \alpha I_4 & + & \beta I_5 & = & R_1 I \\
 & & & & & & I_4 & + & I_5 & = & I
\end{array}$$

mit

$$\alpha = R_{12} + \frac{R_4R_{123}}{R_3} \quad \text{und} \quad \beta = \frac{-R_5R_{123}}{R_3}\,.$$

Schließlich muß noch I_4 aus den letzten beiden Gleichungen eliminiert werden; wir addieren also die vierte Gleichung unverändert zur letzten

und ihr α-faches zur vorletzten:

$$\begin{array}{rcrcrcrcrcr}
I_1 & + & I_2 & & & & & & & = & I \\
 & & I_2 & - & I_3 & - & I_4 & & & = & 0 \\
 & & & & -R_3I_3 & + & R_4I_4 & - & R_5I_5 & = & 0 \\
 & & & & & & -I_4 & + & -I_5 & = & -I \\
 & & & & & & & & (\beta-\alpha)I_5 & = & (R_1-\alpha)I \\
 & & & & & & & & 0 & = & 0
\end{array}$$

Damit können wir nacheinander

$$I_5 = \frac{R_1-\alpha}{\beta-\alpha}\cdot I\,, \qquad I_4 = I - I_5\,, \qquad I_3 = \frac{R_4I_4 - R_5I_5}{R_3}\,,$$
$$I_2 = I_3 + I_4 \qquad \text{und} \qquad I_1 = I - I_2$$

bestimmen, wobei sich der Leser noch überlegen sollte, warum die Division durch $\beta - \alpha$ unproblematisch ist.

Zum Berechnen der Ströme in konkreten Beispielen reicht diese Lösungsformel aus; ist man allerdings an einem symbolischen Ausdruck interessiert, muß man die Definitionen von $\alpha, \beta, R_{12}, R_{123}$ einsetzen und nacheinander alle rechte Seiten auf Ausdrücke nur in I und den R_i reduzieren. Dies ist eine *im Prinzip* einfache Übungsaufgabe in Bruchrechnung, die man allerdings im vorliegenden Fall besser einem Computeralgebrasystem überläßt. Als Ergebnis erhält man die expliziten Formeln

$$I_1 = \frac{\big(R_2R_5 + R_2R_3 + R_2R_4 + R_3R_4\big)\cdot I}{R_1R_5 + R_2R_5 + R_3R_5 + R_1R_3 + R_2R_3 + R_1R_4 + R_2R_4 + R_3R_4}$$
$$I_2 = \frac{\big(R_1R_4 + R_1R_5 + R_3R_5 + R_1R_3\big)\cdot I}{R_1R_5 + R_2R_5 + R_3R_5 + R_1R_3 + R_2R_3 + R_1R_4 + R_2R_4 + R_3R_4}$$
$$I_3 = \frac{\big(R_1R_4 - R_2R_5\big)\cdot I}{R_1R_5 + R_2R_5 + R_3R_5 + R_1R_3 + R_2R_3 + R_1R_4 + R_2R_4 + R_3R_4}$$
$$I_4 = \frac{\big(R_1R_3 + R_1R_5 + R_2R_5 + R_3R_5\big)\cdot I}{R_1R_5 + R_2R_5 + R_3R_5 + R_1R_3 + R_2R_3 + R_1R_4 + R_2R_4 + R_3R_4}$$
$$I_5 = \frac{\big(R_2R_3 + R_1R_4 + R_2R_4 + R_3R_4\big)\cdot I}{R_1R_5 + R_2R_5 + R_3R_5 + R_1R_3 + R_2R_3 + R_1R_4 + R_2R_4 + R_3R_4}\,,$$

die für die meisten *konkreten* Anwendungen erheblich weniger nützlich sind als die obige Form des Ergebnisses.

Als letztes Beispiel schließlich betrachten wir eines, das auch von einem Parameter abhängt, bei dem man aber *nicht* wie im obigen Beispiel einfach durch Ausdrücke im Parameter dividieren darf. (Eigentlich hätten wir es da auch nicht immer dürfen, aber wir haben einfach angenommen, daß alle Widerstände wirklich vorhanden und damit positiv sind; da Kurzschlüsse immer wieder vorkommen, ist diese Annahme nicht hundertprozentig realistisch.) Das Gleichungssystem hänge ab von einem Paramter $a \in k$ und habe die Form

$$\begin{aligned} x_1 + ax_2 + x_3 &= 1 \\ x_1 + x_2 &= 1 \\ -2x_1 - 2ax_2 - ax_3 &= 1\,. \end{aligned}$$

Ein Computeralgebrasystem findet unschwer die Lösung

$$x_1 = \frac{a^2 - 3a - 1}{a^2 - 3a + 2}, \qquad x_2 = \frac{3}{a^2 - 3a + 2}, \qquad x_3 = \frac{-3}{a - 2}\,.$$

Diese „Lösung“ hat aber für $a = 2$ und auch für $a = 1$ Nullen im Nenner, ist dort also nicht erklärt. Für ein Computeralgebrasystem ist das kein Problem: Wie der Name sagt, rechnet es *algebraisch*, und da ist a keine reelle Zahl, sondern ein Symbol, das nichts mit irgendwelchen Zahlen zu tun hat. Damit ist $a - 2$ ein formaler Ausdruck, der nie Null sein kann, denn das *Symbol* a ist schließlich verschieden von der *Zahl* zwei.

Dieses Rechnen in sogenannten Funktionenkörpern ist mathematisch problemlos, ist aber nicht das, was in den meisten Anwendungen gefragt ist: Dort steht a im allgemeinen für eine variable Größe, in Abhängigkeit von der das Gleichungssystem gelöst werden soll. Man kann sich beispielsweise vorstellen, daß das Gleichungssystem ein lineares Regelungsproblem beschreibt in Abhängigkeit von steuerbaren Größen x_1, x_2 und x_3, wobei die zu steuernden Größen zu

$$x_1 + ax_2 + x_3, \quad x_1 + x_2 \quad \text{und} \quad -2x_1 - 2ax_2 - ax_3$$

werden. Der Paramter a wäre dann zu interpretieren als eine von außen vorgegebene Umgebungsbedingung (z.B. die Temperatur), und das lineare Gleichungssystem besagt, daß wir das System so regeln wollen, daß die drei steuerbaren Größen allesamt eins werden – unabhängig von der Außentemperatur.

Bei einer solchen Interpretation können wir natürlich *nicht* einfach durch $a - 2$ dividieren; ein Ergebnis, wie $x_3 = -3/(a-2)$ besagt in so einem Fall, daß das Ziel im Falle $a = 2$ nicht erreichbar ist, und das ist ein sehr wichtiges Ergebnis. Bei einem System, das einen Stromkreis beschreibt, könnte das zum Beisiel bedeuten, daß beim Parameterwert $a = 2$ ein Kurzschluß entsteht, so daß dieser Parameterwert unbedingt verhindert werden muß. Deshalb muß man bei einem Gleichungssystem, das ein reales Problem beschreibt, vor jeden Division durch einen parameterabhängigen Ausdruck garantieren, daß dieser Ausdruck von Null verschieden ist, und man muß die Fälle, in denen er Null wird, gesondert diskutieren.

Im vorliegenden Beispiel (einer Vordiplomsaufgabe vom April 1999) führt dies auf folgende Lösung:

Subtraktion der ersten Gleichung von der zweiten sowie Addition der zweifachen ersten Gleichung zur dritten ergibt

$$\begin{aligned} (a-1)x_2 + x_3 &= 0 \\ (2-a)x_3 &= 3\,. \end{aligned}$$

Für $a = 2$ ist die letzte dieser beiden Gleichungen unlösbar, ansonsten ist

$$x_3 = \frac{3}{2-a} \quad \text{falls} \quad a \neq 2\,.$$

Für $a = 1$ wird die vorletzte Gleichung zu $x_3 = 0$, was der schon gefundenen Lösung

$$x_3 = \frac{3}{2-a} = 1$$

widerspricht; auch dann ist also das Gleichungssystem unlösbar. In allen anderen Fällen erhalten wir

$$x_2 = \frac{3}{(a-1)(a-2)} \quad \text{falls} \quad a \neq 1, 2\,.$$

Schließlich läßt sich noch, beispielsweise aus der zweiten Gleichung des ursprünglichen Systems, x_1 berechnen und wir erhalten als Ergebnis:

Die Lösungsmenge ist
$$\mathcal{L} = \left\{ \left(1 - \frac{3}{(a-1)(a-2)},\ \frac{3}{(a-1)(a-2)},\ \frac{3}{a-2} \right) \right\},$$
falls $a \neq 1, 2$, und
$$\mathcal{L} = \emptyset\,, \quad \text{falls } a = 1 \text{ oder } a = 2$$
ist.

In einem solchen Fall wird man durch die Nullstellen des Nenners gewarnt, daß hier etwas schiefgehen muß; es gibt aber auch Beispiele, in denen ein Computeralgebrasystem Lösungen schlichtweg „übersieht“: Bertrachten wir etwa das lineare Gleichungssystem
$$\begin{array}{rcrcrcl} x_1 & + & & & 2x_3 & = & 9 \\ 2x_1 & + & 3x_2 & + & x_3 & = & 9 \\ & - & x_2 & + & ax_3 & = & a+2\,. \end{array}$$
Ein Computer findet leicht die Lösung
$$x = 7, \quad y = -2 \quad \text{und} \quad z = 1\,.$$
Der GAUSSalgorithmus führt uns aber über das Gleichungssystem
$$\begin{array}{rcrcrcl} x_1 & + & & & 2x_3 & = & 9 \\ & & 3x_2 & - & 3x_3 & = & -9 \\ & - & x_2 & + & ax_3 & = & a+2 \end{array}$$
oder
$$\begin{array}{rcrcrcl} x_1 & + & & & 2x_3 & = & 9 \\ & & x_2 & - & x_3 & = & -3 \\ & - & x_2 & + & ax_3 & = & a+2 \end{array}$$
auf die Endgestalt
$$\begin{array}{rcrcrcl} x_1 & + & & & 2x_3 & = & 9 \\ & & x_2 & - & x_3 & = & -3 \\ & & & & (a-1)x_3 & = & a-1\,, \end{array}$$
in der man nur für $a \neq 1$ aus der letzten Gleichung schließen darf, daß $z = 1$ ist; für $a = 1$ haben wir die immer erfüllte Gleichung $0z = 0$. Die Lösungsmenge ist hier also
$$\mathcal{L} = \{(7, -2, 1)\} \quad \text{für } a \neq 1$$

und
$$\mathcal{L} = \{(-2\lambda + 9, \lambda - 3, \lambda) \mid \lambda \in \mathbb{R}\} \quad \text{für } a = 1\,.$$

Jemand, der Maple hinreichend gut kennt, hätte natürlich auch diese vollständige Lösung damit ermitteln können, aber der einfachstmögliche Befehl reicht definitiv nicht aus – zumindest ein Grund, warum man auch heute noch lernen muß, lineare Gleichungssysteme von Hand zu lösen.

Ein anderer Grund, warum speziell Technische Informatiker das lernen müssen, liegt in der Natur vieler Anwendungen: Lineare Gleichungssysteme müssen beispielsweise gelöst werden bei Steuerungs- und Regelungsproblemen. In vielen Fällen wird diese Steuerung nicht von einem leistungsfähigen Universalrechner durchgeführt, sondern von einer eigens dafür entwickelten Schaltung, die mit möglichst wenig Aufwand arbeiten soll – sei es aus Kostengründen oder wegen des Raumbedarfs oder der Wärmeentwicklung. In solchen Fällen geht es dann darum, die Lösung möglichst effizient zu ermitteln, und bei der Definition des Wortes „effizient“ können hier durchaus auch nichtmathematische Gesichtspunkte eine Rolle spielen. Daher ist es wichtig, das volle Instrumentarium der Lösungstheorie linearer Gleichungssysteme zu beherrschen, um die jeweils beste Methode implementieren zu können. Deshalb werden wir auch noch Alternativen zum GAUSS-Algorithmus betrachten, und in der *Numerik I* werden weitere Verfahren folgen.

f) Die Struktur der Lösungsmenge

Nach diesen Beispielen ist es an der Zeit, wieder zu den theoretischen Grundlagen zurückzukehren. Sei also wieder
$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1m}x_m &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2m}x_m &= b_2 \\ \vdots \qquad\qquad &= \vdots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nm}x_m &= b_n\,, \end{aligned}$$
ein allgemeines lineares Gleichungssystem. Wenn wir die a_{ij} zu einer

Matrix

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

zusammenfassen und die Unbekannten und rechten Seiten zu Vektoren

$$\vec{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_m \end{pmatrix} \quad \text{und} \quad \vec{b} = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix},$$

läßt sich das Gleichungssystem auch kurz als $A\vec{x} = \vec{b}$ schreiben, wobei wir das Produkt der Matrix A mit dem Vektor $\vec{x}$ so definieren, daß es jener Vektor aus k^n sein soll, dessen i-te Komponente die Summe

$$\sum_{j=1}^{m} a_{ij} x_j$$

sein soll. Identifiziert man Vektoren aus k^m mit $m \times 1$-Matrizen, ist das gerade das Produkt der $n \times m$-Matrix A mit der $m \times 1$-Matrix der Variablen.

Wir bezeichen A als die *Matrix des Gleichungssystems* und

$$\left(A \mid \vec{b}\,\right) \underset{\text{def}}{=} \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1m} & b_1 \\ a_{21} & a_{22} & \dots & a_{2m} & b_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nm} & b_m \end{pmatrix}$$

als die *erweiterte Matrix.*

Damit können wir das Gleichungssystem in die Sprache der Vektorräume und linearen Abbildungen übersetzen: Wir betrachten die lineare Abbildung

$$\varphi\colon k^m \to k^n; \quad \vec{v} \mapsto A\vec{v}\,,$$

und die Lösungsmenge des linearen Gleichungssystems ist

$$\mathcal{L} = \left\{\vec{v} \in k^m \;\middle|\; \varphi(\vec{v}) = \vec{b}\right\}.$$

Für zwei Lösungsvektoren $\vec{v}, \vec{w} \in k^m$ ist

$$\varphi(\vec{v} - \vec{w}) = \varphi(\vec{v}) - \varphi(\vec{w}) = \vec{b} - \vec{b} = \vec{0}\,,$$

die Differenz zweier Lösungen ist also eine Lösung des entsprechenden linearen Gleichungssystems mit lauter Nullen auf der rechter Seite.

Definition: Ein lineares Gleichungssystem $A\vec{x} = \vec{b}$ heißt *homogen*, wenn $\vec{b} = \vec{0}$ ist; ansonsten heißt es *inhomogen.* $A\vec{x} = \vec{0}$ heißt das zu $A\vec{x} = \vec{b}$ gehörige homogene Gleichungssystem.

Damit wissen wir also

Lemma: Zwei Lösungen des linearen Gleichungssystems $A\vec{x} = \vec{b}$ unterscheiden sich durch eine Lösung des homogenen Gleichungssystems. ■

Die Lösung eines homogenen Gleichungssystems besteht aus allen Vektoren aus k^m, die von der oben definierten linearen Abbildung φ auf den Nullvektor abgebildet werden, ist also gerade der Kern von φ und damit ein Untervektorraum von k^n. Insbesondere ist die Lösungsmenge eines homogenen linearen Gleichungssystems also nie leer: Wie man sofort auch direkt sieht, gibt es immer die sogenannte *triviale* Lösung, bei der alle Variablen gleich Null sind.

Nach der Dimensionsformel aus §1*i)* ist

$$\dim \operatorname{Kern} \varphi = m - \dim \operatorname{Bild} \varphi\,,$$

und das Bild wird erzeugt von den m Vektoren $\varphi(\vec{e}_i) = A\vec{e}_i$. Das sind aber gerade die Spaltenvektoren der Matrix A; die Dimension des Bildes ist also gleich der Anzahl linear unabhängiger Spaltenvektoren von A, und diese Zahl hatten wir oben als den (Spalten-)Rang der Matrix definiert. Damit wissen wir also

Lemma: Die Lösungsmenge des homogenen linearen Gleichungssystems $A\vec{x} = \vec{0}$ in m Variablen ist ein Untervektorraum der Dimension $m - \operatorname{Rang} A$. ■

Im Falle eines inhomogenen Gleichungssystems ist die Sache nicht ganz so einfach: Schließlich hatten wir bei den Beispielen schon gesehen, daß von zwei Gleichungssystemen, die sich nur in ihrer rechten Seite unterscheiden, das eine unlösbar sein kann, während das andere eine oder mehrere Lösungen hat.

Der Grund dafür wird klar, wenn wir das Gleichungssystem wieder über die lineare Abbildung φ interpretieren: Das Gleichungssystem $A\vec{x} = \vec{b}$ ist genau dann lösbar, wenn die rechte Seite $\vec{b}$ im Bild von φ liegt. Damit muß sie linear abhängig sein von den Spaltenvektoren von A, die ja den Bildraum erzeugen, d.h. der Rang der *erweiterten* Matrix, die außer den Spaltenvektoren von A noch den Vektor $\vec{b}$ enthält, darf nicht größer sein, als der von A. Kleiner kann er natürlich nicht sein, also haben wir gezeigt

Lemma: Ein inhomogenes Gleichungssystem $A\vec{x} = \vec{b}$ ist genau dann lösbar, wenn der Rang r seiner erweiterten Matrix gleich dem von A ist. Es ist genau dann eindeutig lösbar, wenn dieser gemeinsame Rang gleich der Anzahl m der Variablen ist. ■

Ist der gemeinsame Rang von Matrix und erweiterter Matrix kleiner als m, so ist das Gleichungssystem nicht eindeutig lösbar. Da sich zwei Lösungen um eine Lösung des zugehörigen homogenen Gleichungssystems unterscheiden und diese wiederum einen Vektorraum der Dimension $m - \text{Rang}\, A$ bilden, hängt die Lösung dann von $m - \text{Rang}\, A$ Parametern ab, ist aber für ein echt inhomogenes Gleichungssystem kein Vektorraum: Ein Vektorraum enthält stets den Nullvektor, und wenn es eine Lösung gibt, bei der *alle* Variablen verschwinden, stehen auf der rechten Seite des Gleichungssystems lauter Nullen, d.h. das Gleichungssystem ist homogen.

Das obige Lemma wird nur selten von praktischem Nutzen sein, denn wenn Matrix und erweiterte Matrix keine sehr spezielle Gestalt haben, wird der GAUSS-Algorithmus im allgemeinen die einfachste Möglichkeit sein, um die Ränge von Matrix und erweiterter Matrix zu bestimmen. Man muß ihn allerdings nicht ganz zu Ende durchführen, denn die Ränge sind schon klar, wenn auf der linken Seite Trepengestalt erreicht ist. Der Vollständigkeit halber sei hier kurz angegeben, wie diese im allerallgemeinsten Fall aussieht:

Die verschiedenen Schritte des GAUSS-Algorithmus wirken sich auf die Matrix wie auch auf die erweiterte Matrix so aus, daß entweder zwei Zeilen miteinander vertauscht werden, oder aber ein Vielfaches einer Zeile zu einer anderen Zeile addiert wird – Addition bedeutet dabei die gewöhnliche komponentenweise Addition, d.h. die Addition von Zeilenvektoren. Am Ende entsteht eine erweiterte Matrix der Form

$$\left(\begin{array}{ccccccccccc}
[0] & \bullet & [\star] & \star & [\star] & \star & \dots & [\star] & \star & [\star] & \star \\
[0] & 0 & [0] & \bullet & [\star] & \star & \dots & [\star] & \star & [\star] & \star \\
[0] & 0 & [0] & 0 & [0] & \bullet & \dots & [\star] & \star & [\star] & \star \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots \\
[0] & 0 & [0] & 0 & [0] & 0 & \dots & [0] & \bullet & [\star] & \star \\
\hline
[0] & 0 & [0] & 0 & [0] & 0 & \dots & [0] & 0 & [0] & \bullet \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots \\
[0] & 0 & [0] & 0 & [0] & 0 & \dots & [0] & 0 & [0] & \bullet \\
\hline
[0] & 0 & [0] & 0 & [0] & 0 & \dots & [0] & 0 & [0] & 0 \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots \\
[0] & 0 & [0] & 0 & [0] & 0 & \dots & [0] & 0 & [0] & 0
\end{array}\right),$$

wobei $\bullet$ für ein von Null verschiedenes und $\star$ für ein beliebiges Element des Körpers k steht; $[0]$ und $[\star]$ stehen für keine, eine oder mehrere Nullen *bzw.* beliebige Elemente, deren Anzahl in untereinanderstehenden Termen jeweils die gleiche sein soll.

Die Zeilen zwischen den beiden eingezeichneten Linien, in denen alle Einträge der Matrix des Gleichungssystems verschwinden, müssen natürlich nicht wirklich auftreten, genauso wenig die Zeilen unterhalb der zweiten Linie, wo sogar alle Einträge der erweiterten Matrix verschwinden, so daß auch die rechten Seiten der entsprechenden Gleichungen in der Endgestalt nach Anwendung des GAUSS-Algorithmus Null sind.

Falls zwischen den beiden eingezeichneten Linien wirklich eine oder mehrere Zeilen stehen, ist das Gleichungssystem *unlösbar,* denn dann gibt es ja in der Endform des Gleichungssystems Gleichungen, in denen links alle Koeffizienten Null sind, während rechts eine von Null verschiedene Zahl steht. Indem man gegebenenfalls ein Vielfaches der ersten solchen Gleichung von den etwa vorhandenen weiteren Gleichungen subtrahiert, kann man erreichen, daß alle weiteren Gleichungen zu 0 = 0 werden, d.h. wir können erreichen, daß zwischen den beiden Linien *höchstens eine* Zeile steht, und das lineare Gleichungssystem ist genau dann lösbar, wenn keine dort steht.

Damit ist klar, daß die Anzahl der Zeilen oberhalb der ersten Linie der *Rang der Matrix A* ist und die Anzahl der Zeilen oberhalb der zweiten Linie (nachdem man dafür Sorge getragen hat, daß höchstens eine Zeile zwischen den beiden Linien steht) der *Rang der erweiterten Matrix.*

Noch etwas läßt sich diesem Diagramm ansehen: Der Rang der Matrix, die maximale Anzahl linear unabhängiger Spalten also, ist offenbar gerade gleich der Anzahl der Zeichen • oberhalb des ersten Strichs. Genau das ist aber auch die maximale Anzahl linear unabhängiger Zeilen der Matrix, d.h. wir können am Rande noch notieren, daß der *Zeilrang* einer Matrix gleich dem *Spaltenrang* ist, was die Kurzbezeichnung *Rang* rechtfertigt.

Betrachten wir zum Abschluß noch ein Beispiel für das Rangkriterium zur Lösbarkeit eines linearen Gleichungssystems: Im vorigen Abschnitt hatten wir das lineare Gleichungssystem

$$\begin{array}{rcrcrcrcrcrcr}
 & & x_2 & & & + & 5x_4 & - & 2x_5 & + & x_6 & = & 10\\
2x_1 & + & 2x_2 & & & - & 4x_4 & - & 6x_5 & + & 2x_6 & = & 12\\
6x_1 & - & 3x_2 & + & x_3 & + & 5x_4 & & & - & 3x_6 & = & 15\\
4x_1 & - & 7x_2 & & & - & 3x_4 & & & + & x_6 & = & 0\\
2x_1 & - & 7x_2 & + & x_3 & + & 13x_4 & + & 12x_5 & - & 7x_6 & = & -8
\end{array}$$

als unlösbar erkannt. Seine erweiterte Matrix ist

$$\begin{pmatrix}
0 & 1 & 0 & 5 & -2 & 1 & 10\\
2 & 2 & 0 & -4 & -6 & 2 & 12\\
6 & -3 & 1 & 5 & 0 & -3 & 15\\
4 & -7 & 0 & -3 & 0 & 1 & 0\\
2 & -7 & 1 & 13 & 12 & -7 & -8
\end{pmatrix},$$

und nach Anwendung des GAUSS-Algorithmus kommen wir auf die Endgestalt

$$\left(\begin{array}{cccccc|c}
2 & 2 & 0 & -4 & -6 & 2 & 12\\
0 & 1 & 0 & 5 & -2 & 1 & 10\\
0 & 0 & 1 & 62 & 0 & 0 & 69\\
0 & 0 & 0 & 20 & 6 & 0 & 6\\
\hline
0 & 0 & 0 & 0 & 0 & 0 & 1\\
\hline
\end{array}\right).$$

Somit hat die Matrix des linearen Gleichungssystems den Rang vier, wohingegen die erweiterte Matrix den Rang fünf hat.

Ändern wir im Gleichungssystem die rechte Seite der letzten Gleichung von -8 zu -9, so wird es lösbar. In der Endgestalt ändert sich, wie wir oben gesehen haben, auch nur die rechte Seite der letzten Gleichung; die erweiterte Matrix wird also schließlich zu

$$\left(\begin{array}{cccccc|c}
2 & 2 & 0 & -4 & -6 & 2 & 12\\
0 & 1 & 0 & 5 & -2 & 1 & 10\\
0 & 0 & 1 & 62 & 0 & 0 & 69\\
0 & 0 & 0 & 20 & 6 & 0 & 6\\
\hline\hline
0 & 0 & 0 & 0 & 0 & 0 & 0
\end{array}\right).$$

Hier haben Matrix und erweiterte Matrix denselben Rang vier.

g) Affine Räume

Für inhomogene Gleichungssysteme kennen wir bislang noch nicht die Struktur des Lösungsraums selbst; wir wissen nur, daß die *Differenzen* von Lösungen (so es welche gibt) einen Vektorraum bilden.

Dies sollte etwas an die Situation in §1*a)* erinnern, wo wir zwei Punkten deren *Verbindungsvektor* zugeordnet hatten und zwei Vektoren unabhängig von ihren Anfangs- und Endpunkten als gleich bezeichnet hatten, wenn sie nur die gleiche Länge und Richtung hatten. Eine ähnliche Situation haben wir auch hier: Lösungen eines inhomogenen Gleichungssystems (so es welche gibt) verhalten sich wie Punkte, Lösungen eines inhomogenen Gleichungssystems wie Vektoren.

Beispielsweise gibt es keine sinnvolle Addition zweier Punkte, genauso wie auch die Summe zweier Lösungen einen inhomogenen Gleichungssystems keine Lösung mehr ist. Andererseits kann man einen Vektor an