Lineare Algebra und Geometrie

Ulrich Bunke*

3. Dezember 2004

Inhaltsverzeichnis

*Göttingen, bunke@uni-math.gwdg.de

1 Zahlen und Systeme von Gleichungen

1.1 Zahlen

1.1.1 Natürliche Zahlen, Induktionsbeweise, Anzahlbegriff für endliche Mengen

Wir bezeichnen mit $\mathbb{N}$ die Menge der natürlichen Zahlen. Wie üblich schreiben wir

$$\mathbb{N} = \{1, 2, \dots\} .$$

Wenn wir die Null mit einschließen wollen, dann schreiben wir

$$\mathbb{N}_0 := \mathbb{N} \cup \{0\} .$$

Natürliche Zahlen kann man addieren und multiplizieren:

$$+ : \mathbb{N} \times \mathbb{N} \to \mathbb{N} \quad , \quad (a,b) \mapsto a+b$$
$$* : \mathbb{N} \times \mathbb{N} \to \mathbb{N} \quad , \quad (a,b) \mapsto a*b .$$

Oft läßt man das Symbol für die Multiplikation weg und schreibt ab für $a*b$. Diese Operationen sind kommutativ, d.h. sie erfüllen die Relationen

$$a+b = b+a, \quad a*b = b*a ,$$

und sie erfüllen das Distributivgesetz, d.h.

$$a*(b+c) = a*b+a*c .$$

Genauer müsste man für die Kommutativität der Addition so formulieren:

$$\textit{Für alle } a,b \in \mathbb{N} \textit{ gilt die Gleichung } a+b = b+a \quad .$$

Entsprechendes gilt für die anderen beiden Identitäten.

Die obigen Eigenschaften charakterisieren $(\mathbb{N}, +, *)$ als einen Halbring.

Eine weiter wesentliche Eigenschaft der natürlichen Zahlen ist die Ordnung. Wir verstehen diese als eine Relation, also als eine Abbildung

$$\leq : \mathbb{N} \times \mathbb{N} \to \{wahr, falsch\} .$$

Wir schreiben üblicherweise $\leq (a,b) =: a \leq b$. So gilt etwa

$$5 \leq 7 = wahr , \quad 6 \leq 3 = falsch .$$

Eine Ordnungsrelation $\leq$ auf einer Menge A muß folgende Eigenschaften haben:

1. (transitiv) Die Relationen $a \leq b$ und $b \leq c$ implizieren $a \leq c$.

2. (reflexiv) Es gilt $a \leq a$.

3. (antisymmetrisch) Falls $a \neq b$ ist, gilt entweder $a \leq b$ oder $b \leq a$.

Die Ordnung der natürlichen Zahlen hat diese Eigenschaften.

Jede Teilmenge der natürlichen Zahlen hat ein kleinstes Element. Für $n \in \mathbb{N}$ betrachten wir die Teilmenge $\mathbb{N}^{<n} := \{m \in \mathbb{N} | n < m\} \subset \mathbb{N}$. Das kleinste Element von $\mathbb{N}^{<n}$ heißt Nachfolger von n und ergibt sich durch $n+1$. Jede natürliche Zahl hat also einen Nachfolger. Es gilt folgende wichtige Bemerkung.

Sei $A \subset \mathbb{N}$ eine Teilmenge mit folgenden Eigenschaften:

1. $1 \in A$

2. *Für jedes $a \in A$ ist auch der Nachfolger $a+1 \in A$.*

Dann gilt $A = \mathbb{N}$. Diese Beobachtung ist die Basis des Beweisprinzips der vollständigen Induktion.

Als Beispiel wollen wir die Formel

$$\sum_{k=1}^{n} k = \frac{n(n+1)}{2}$$

zeigen. Wir formulieren dies wie folgt um. *Für jede natürliche Zahl $n \in \mathbb{N}$ gilt $\sum_{k=1}^{n} k = \frac{n(n+1)}{2}$.* Nun zum Beweis. Sei $A \subset \mathbb{N}$ die Teilmenge derjenigen natürlichen Zahlen, für welche die Identität $\sum_{k=1}^{n} k = \frac{n(n+1)}{2}$ gilt. Wir zeigen den sogenannten Induktionsanfang: *Es gilt $1 \in A$:* In der Tat gilt $\sum_{k=1}^{1} k = 1$ und $\frac{1(1+1)}{2} = 1$. Nun kommen wir zum sogenannten Induktionsschritt. *Wir zeigen, daß mit $n \in A$ auch der Nachfolger $n+1 \in A$ ist.* In der Tat, sei $n \in A$. Dann gilt $\sum_{k=1}^{n} k = \frac{n(n+1)}{2}$. Dann rechnen wir

$$\sum_{k=1}^{n+1} k = \sum_{k=1}^{n} k + (n+1) = \frac{n(n+1)}{2} + (n+1) = \frac{n(n+1)+2(n+1)}{2} = \frac{(n+1)((n+1)+1)}{2}.$$

Dies besagt aber $n+1 \in A$: Die gezeigten Aussagen implizieren, daß $A = \mathbb{N}$ gelten muß, womit der Beweis beendet ist.

Die natürlichen Zahlen benutzt man, um die Anzahl der Elemente in endlichen Mengen zu beschreiben. Dazu müssen wir zuerst einmal endliche Mengen von unendlichen unterscheiden. Die Anzahl der Elemente verschiedener Mengen (deren Mächtigkeit) vergleicht man mit Hilfe von Abbildungen. Genauer:

Definition 1.1. *Zwei Mengen A, B heißen gleichmächtig, wenn es eine bijektive Abbildung $A \xrightarrow{\sim} B$ gibt.*

"Gleichmächtigkeit" ist eine Äquivalenzrelation für Mengen. Sei $\sim$ eine Relation auf A, also eine Abbildung

$$\sim: A \times A \to \{wahr, falsch\} .$$

Wir zögern hier, A eine Menge zu nennen, weil wir gleich für A die Gesamtheit aller Mengen nehmen wollen, welches keine Menge ist. Wir schreiben wieder $\sim (a,b) =: a \sim b$. Eine Äquivalenzrelation erfüllt

1. (reflexiv) Es gilt $a \sim a$.
2. (transitiv) Die Relationen $a \sim b$ und $b \sim c$ implizieren $a \sim c$.
3. (symmetrisch) Es gilt $a \sim b$ genau dann wenn $b \sim a$.

In der Tat hat "Gleichmächtigkeit" diese Eigenschaften. Es gilt $M \sim M$ wegen der Existenz der Bijektion $\mathtt{id}_M : M \to M$. Die Relationen $M \sim N$ und $N \sim P$ implizieren $M \sim P$. In der Tat, seien $f : M \to N$ und $g : N \to P$ Bijektionen, dann ist $g \circ f : M \to P$ auch eine Bijektion. Wenn $M \sim N$ gilt, dann auch $N \sim M$. In der Tat, wenn $f : M \to N$ eine Bijektion ist, dann ist die inverse Abbildung $f^{-1} : N \to M$ eine Bijektion.

Wir kommen nun zur Charakterisierung endlicher Mengen.

Definition 1.2. *Eine Menge A heißt endlich, wenn es keine echte Teilmenge $B \subset A$ gibt, welche zu A gleichmächtig ist. Andernfalls heißt A unendlich.*

Die Menge $\mathbb{N}$ ist unendlich. In der Tat ist $\mathbb{N} \setminus \{1\} \subset \mathbb{N}$ eine zu $\mathbb{N}$ gleichmächtige echte Teilmenge. Die geforderte Bijektion kann etwa durch $x \mapsto x - 1$ gegeben werden.

Lemma 1.3. *Es gibt genau eine Zuordnung $A \to |A|$ welche jeder endlichen Menge A eine Zahl $|A| \in \mathbb{N}_0$ zuweist, so daß für disjunkte A, B die Gleichung*

$$|A \cup B| = |A| + |B|$$

gilt. Diese Zuordnung erfüllt weiterhin

$$|A \times B| = |A| \times |B| \ , \quad |A \cup B| + |A \cap B| = |A| + |B| \ .$$

Wir beweisen dieses Lemma hier nicht. Die Argumentation hängt nämlich davon ab, wie man $\mathbb{N}$ definiert. Versteht man $\mathbb{N}$ als die Menge der Äquivalenzklassen endlicher Mengen bezüglich der Relation "gleichmächtig", dann ist der erste Teil des Lemmas tautologisch.

Mit den natürlichen Zahlen werden in der Praxis jede Art von Zählproblemen modelliert.

1.1.2 Die ganzen Zahlen

Mit $\mathbb{Z}$ bezeichnen wir die ganzen Zahlen. Wir schreiben

$$\mathbb{Z} = \{\ldots, -2, -1, 0, 1, 2, \ldots\} \ .$$

Ganze Zahlen kann man addieren und multiplizieren:

$$\begin{aligned} + &: \mathbb{Z} \times \mathbb{Z} \to \mathbb{Z} \quad , \quad (a,b) \mapsto a + b \\ * &: \mathbb{Z} \times \mathbb{Z} \to \mathbb{Z} \quad , \quad (a,b) \mapsto a * b \ . \end{aligned}$$

Diese Operationen sind kommutativ und distributiv. Zusätzlich kann man auch subtrahieren, d.h. die Gleichung $a + x = b$ hat für alle ganzen Zahlen a, b eine eindeutige Lösung $x = b - a$. Es gilt weiterhin die Relation $a + 0 = a$. Mit diesen Eigenschaften wird $(\mathbb{Z}, +, *, 0, 1)$ als ein Ring charakterisiert.

Wir können die ganzen Zahlen als formale Differenzen $m - n$ von natürlichen Zahlen $m, n \in \mathbb{N}$ verstehen. Dabei ist zu beachten, daß $(m + a) - (n + a)$ die gleiche ganze Zahl wie $m - n$ darstellt. Die natürlichen Zahlen können als Teilmenge der ganzen Zahlen verstanden werden, $\mathbb{N} \subset \mathbb{Z}$.

Die ganzen Zahlen werden zur Modellierung von Abzählproblemen benutzt, in welchen auch Schulden (negative Anzahlen) erlaubt sind. Beispiele sind etwa Kontostände (in Cent) oder elektrische Ladungen.

1.1.3 Die rationalen Zahlen

Mit $\mathbb{Q}$ bezeichnen wir die Menge der rationalen Zahlen. Rationale Zahlen kann man addieren und multiplizieren:

$$+ : \mathbb{Q} \times \mathbb{Q} \to \mathbb{Q} \quad , \quad (a,b) \mapsto a+b$$
$$* : \mathbb{Q} \times \mathbb{Q} \to \mathbb{Q} \quad , \quad (a,b) \mapsto a*b \; .$$

Diese Operationen sind kommutativ und distributiv. Man kann wie bei den ganzen Zahlen subtrahieren. Zusätzlich kann man durch von Null verschiedene Elemente dividieren, d.h. die Gleichung $a*x=b$ hat für alle rationalen Zahlen a,b mit $a \neq 0$ eine eindeutige Lösung $x = \frac{b}{a}$. Es gilt weiterhin die Relation $a*0=0$ und $a*1=a$. Mit diesen Eigenschaften wird $(\mathbb{Q},+,*,0,1)$ als ein Körper charakterisiert.

Wir können die rationalen Zahlen als formale Brüche $\frac{m}{n}$ von ganzen Zahlen $m,n \in \mathbb{Z}$ mit $n \neq 0$ verstehen. Dabei ist zu beachten, daß für $a \neq 0$ der Bruch $\frac{a*m}{a*n}$ die gleiche ganze Zahl wie $\frac{m}{n}$ darstellt. Die ganzen Zahlen können als Teilmenge der rationalen Zahlen verstanden werden, $\mathbb{Z} \subset \mathbb{Q}$.

In der Praxis benutzt man rationale Zahlen, um Größen zu modellieren, welche durch Aufteilung von Ganzheiten entstehen. Ein typisches Beispiel ist die Bildung von Notendurchschnitten. Die wesentliche Bedeutung der rationalen Zahlen ist jedoch die als der kleinste Körper, welcher den Ring der ganzen Zahlen enthält.

1.1.4 Die reellen Zahlen

Mit $\mathbb{R}$ bezeichnen wir die Menge der rellen Zahlen. Reelle Zahlen kann man addieren und multiplizieren, wobei die Körpereigenschaften wie bei den rationalen Zahlen erfüllt sind. Die reellen Zahlen enthalten die rationalen Zahlen, $\mathbb{Q} \subset \mathbb{R}$.

Eine wesentliche Eigenschaft der reellen Zahlen (und ihrer Teilmengen $\mathbb{N} \subset \mathbb{Z} \subset \mathbb{Q} \subset \mathbb{R}$) ist die Anordnung, daß heißt, auf $\mathbb{R}$ ist eine Ordnungsrelation "$\leq$" definiert. Die Verträglichkeit dieser Relation mit den Körperstrukturen wird in der Analysis beschrieben. Mit Hilfe der Anordnung kann man (abgeschlossene) Intervalle

$$[a,b] := \{x \in \mathbb{R} | a \leq x \leq b\} \subset \mathbb{R}$$

definieren. Die reellen Zahlen sind vollständig, d.h. für jede Folge von geschachtelten (abgeschlossenen) Intervallen

$$\ldots I_n \subset I_{n-1} \subset \cdots \subset I_0$$

ist der Durchschnitt

$$\bigcap_{n\in\mathbb{N}} I_n$$

nicht leer. Man kann $\mathbb{R}$ als den kleinsten vollständigen $\mathbb{Q}$-enthaltenden Körper charakterisieren.

Eine wesentliche Folge der Anordnung und Vollständigkeit ist die Existenz von Potenzen $a^b \in \mathbb{R}$ für $a > 0$ und $b \in \mathbb{R}$. Dies wird in der Analysis abgeleitet. Insbesondere existieren in $\mathbb{R}$ die Wurzeln $a^{\frac{1}{n}}$ aus positiven Zahlen.

Die meisten "kontinuierlichen" physikalischen Größen werden durch reelle Zahlen modelliert. Typische Beispiele sind Längen, Zeiten, Massen, Kräfte, Geschwindigkeiten, Spannungen, Ströme.

1.2 Lineare Gleichungen mit mehreren Unbekannten

1.2.1 Mehrdimensionale Größen

Mehrdimensionale Größen benutzt man, um komplexe reale Systeme zu modellieren.

Im folgenden Beispiel mögen Kontostände durch reelle Zahlen modelliert werden (dies bedeutet in der Realität, daß bis auf einen sehr kleinen Bruchteil eines Cent gerechnet wird).

Wir denken uns jetzt eine Situation mit mehreren (etwa k-Stück) Konten zu unterschiedlichen Konditionen (das i-te Konto habe etwa den Zinssatz $a_i \in \mathbb{R}$).

Das Geldvermögen wird dann durch die gleichzeitige Angabe aller Kontostände beschrieben. Sei $z_i \in \mathbb{R}$ der Stand des i-ten Kontos.

Sei

$$\mathbb{R}^k := \underbrace{\mathbb{R} \times \cdots \times \mathbb{R}}_{k\times} = \{(x_1, \ldots, x_k) | x_i \in \mathbb{R}\}$$

die Menge der (geordneten) k-Tupel reeller Zahlen. In der Tat, um mit späteren Konventionen verträglich zu sein, werden wir die Elemente als Spalten schreiben:

$$x = (x_1, \ldots, x_k)^t = \begin{pmatrix} x_1 \\ \vdots \\ x_k \end{pmatrix} .$$

Der obere Index "t" bedeutet also, daß die Zeile als Spalte betrachtet wird (wir schreiben Zeilen aus Papierspargründen).

Das Geldvermögen wird also durch einen Punkt

$$z := (z_1, \ldots, z_k)^t \in \mathbb{R}^k$$

beschrieben.

Wenn ich eine Geldmenge u einzahlen will, dann kann ich sie auf die einzelnen Konten aufsplitten. Die Einzahlung auf das i-te Konto ist x_i, und es muß $\sum_{i=1}^k x_i = u$ gelten. Insbesondere kann ich die Einzahlung auch durch einen Punkt

$$x := (x_1, \ldots, x_k)^t \in \mathbb{R}^k$$

beschreiben. Der neue i-te Kontostand nach erfolgter Einzahlung ist $z_i + x_i$. Wir definieren eine Addition

$$+ : \mathbb{R}^k \times \mathbb{R}^k \to \mathbb{R}^k$$

durch komponentenweise Addition:

$$(x_1, \ldots, x_k)^t + (y_1, \ldots, y_k)^t := (x_1 + y_1, \ldots, x_k + y_k)^t .$$

Dann kann ich das Vermögen nach erfolgter Einzahlung als den Punkt $z + x$ beschreiben.

Nun kommt es zu einer Währungsumstellung durch welche die Zahlenbeträge um den Umrechnungsfaktor $\lambda \in \mathbb{R}$ abgeändert werden. Diese betrifft alle Konten in gleicher Weise. Nach der Umstellung ist der Kontostand $(\lambda z_1, \ldots, \lambda z_k)^t$. Es ist sinnvoll, auf $\mathbb{R}^k$ die (skalare) Multiplikation mit den reellen Zahlen

$$* : \mathbb{R} \times \mathbb{R}^k \to \mathbb{R}^k$$

komponentenweise zu definieren:

$$\lambda(x_1, \ldots, x_k)^t := (\lambda x_1, \ldots, \lambda x_k)^t .$$

Der Raum $\mathbb{R}^k$ mit der beschriebenen Addition und der skalaren Multiplikation ist unser Prototyp eines Vektorraumes über $\mathbb{R}$.

Wir betrachten nun den Zinsertrag im Jahr (der Kontostand möge konstant gewesen sein). Das i-te Konto wirft im Jahr ein Zinseinkommen von $z_i a_i$ ab. Folglich belaufen sich meine Zinseinahmen im Jahr auf

$$a_1 z_1 + \cdots + a_k z_k = \sum_{i=1}^{k} a_i z_i \ .$$

Dies kann mit einigen Verabredungen kürzer geschrieben werden. Wir betrachten eine weitere Kopie $\hat{\mathbb{R}}^k$ der Menge $\mathbb{R}^k$, deren Elemente wir jetzt jedoch als Zeilen schreiben werden. Wir definieren weiter die Paarung

$$\hat{\mathbb{R}}^k \times \mathbb{R}^k \to \mathbb{R} \ , \quad (x,y) \mapsto x \bullet y := \sum_{i=1}^{k} x_i y_i \ .$$

Betrachtet man $a = (a_1, \ldots, a_k) \in \hat{\mathbb{R}}^k$, dann sind die Zinseinkünfte durch $a \bullet z$ beschrieben. Die Paarung hat folgende Eigenschaften:

$$\begin{aligned} a \bullet (x+y) &= a \bullet x + a \bullet y \\ (a+b) \bullet x &= a \bullet x + b \bullet x \\ (\lambda a) \bullet x &= \lambda (a \bullet x) \\ a \bullet (\lambda x) &= \lambda (a \bullet x) \end{aligned} \tag{1}$$

Genauer müßte man etwa bei der letzten Identität sagen: *Für alle* $a \in \hat{\mathbb{R}}^k$, $x \in \mathbb{R}^k$ *und* $\lambda \in \mathbb{R}$ *gilt* $a \bullet (\lambda x) = \lambda (a \bullet x)$. Die letzte Regel besagt zum Beispiel, daß es egal ist, ob erst die Währung umgestellt oder erst die Zinsen berechnet werden.

1.2.2 Systeme mit zwei Unbekannten

Wir betrachten zwei leere Konten. Die Zinssätze seien durch $a = (a_1, a_2) \in \hat{\mathbb{R}}^2$ gegeben. Wir möchten die Summe u so einzahlen, daß nach einem Jahr der vorgegebene Betrag e als Zinsen eingenommen wird.

Man muß also die Summe in noch unbekannter Weise aufsplitten:

$$z_1 + z_2 = u \ .$$

Die Zinseinnahmen sind dann

$$a_1 z_1 + a_2 z_2 = e \ .$$

Daraus kann man z_i bestimmen. In diesem einfachen Fall geht das durch schrittweise Elimination.

Wir setzen $z_1 = u - z_2$ in die zweite Gleichung ein. Es ergibt sich

$$(a_2 - a_1) z_2 = e - a_1 u \ .$$

Also gilt (falls $a_1 \neq a_2$)

$$z_2 = \frac{e - a_1 u}{a_2 - a_1}, \quad z_1 = u - z_2 \ .$$

Falls $a_1 = a_2$ und $e = a_1 u$ gilt, dann ist es egal, wie man die Summe aufspaltet. Falls $a_1 = a_2$ ist und $e \neq a_1 u$, dann hat das System keine Lösung.

Wir sehen also, daß ein System aus zwei linearen Gleichungen ein sehr unterschiedliches Lösungsverhalten haben kann, etwa eine eindeutige Lösung, viele Lösungen, oder gar keine Lösung. Die allgemeine Lösungstheorie werden wir später betrachten. Im Beispiel mag noch interessant sein, ob etwa die z_i nicht-negativ sind. Solche Fragen sind nicht Bestandteil der allgemeinen Theorie.

Mit Hilfe des $\bullet$ können wir das Gleichungssystem noch kürzer schreiben. Wir betrachten das Element $b = (1,1) \in \hat{\mathbb{R}}^2$. Dann kann man die beiden Gleichungen kurz als

$$\begin{aligned} a \bullet z &= e \\ b \bullet z &= u \end{aligned} \tag{2}$$

schreiben. Mit Hilfe von Matrizen geht es noch kürzer.

1.2.3 Matrizen

Sei I eine Menge. Dann kann man die Menge der Abbildungen $\mathbb{R}^I = \{x : I \to \mathbb{R}\}$ betrachten. Ein Element aus $\mathbb{R}^I$ kann als ein durch I-indiziertes Tupel $(x_i)_{i \in I}$ aus reellen Zahlen verstanden werden, wenn man $x_i := x(i)$ setzt.

Insbesondere können wir für $n \in \mathbb{N}$ die Menge $\bar{n} := \{1, \ldots, n\}$ betrachten. Offensichtlich kann man $\mathbb{R}^{\bar{n}}$ mit $\mathbb{R}^n$ identifizieren.

Die Elemente der Menge $\bar{n} \times \bar{k}$ sind Paare (i, j) natürlicher Zahlen mit $1 \leq i \leq n$ und $1 \leq j \leq k$. Diese ordnet man der Anschaulichkeit halber in der folgenden Weise an:

$$\begin{pmatrix} (1,1) & \dots & (1,k) \\ \vdots & & \vdots \\ (n,1) & \dots & (n,k) \end{pmatrix}$$

Definition 1.4. *Die Menge der reell-wertigen $n \times k$-Matrizen wird durch*

$$\mathtt{Mat}(n,k,\mathbb{R}) := \mathbb{R}^{\bar{n} \times \bar{k}}$$

gegeben.

Folglich ist eine $n \times k$-Matrix M durch eine Familie von Zahlen $(M_{i,j})_{1 \leq i \leq n, 1 \leq j \leq k}$ gegeben. Um eine solche Matrix aufzuschreiben, ersetzt man in der obigen Anordnung das Paar (i, j) durch den Wert $M_{i,j}$:

$$M = \begin{pmatrix} M_{1,1} & \dots & M_{1,k} \\ \vdots & & \vdots \\ M_{n,1} & \dots & M_{n,k} \end{pmatrix}$$

Man kann die Zeilen von M als Elemente aus $\hat{\mathbb{R}}^k$ verstehen. Dann kann man M als n-Spalte schreiben:

$$M = \begin{pmatrix} \hat{m}_1 \\ \vdots \\ \hat{m}_n \end{pmatrix} , \quad \hat{m}_i = (M_{i,1}, \dots, M_{i,k}) .$$

Analog kann man die Spalten von M als Elemente aus $\mathbb{R}^n$ auffassen:

$$M = (m_1, \dots, m_k) , \quad m_i = \begin{pmatrix} M_{1,i} \\ \vdots \\ M_{n,i} \end{pmatrix} .$$

Insbesondere gilt $\mathbb{R}^k = \mathtt{Mat}(k,1,\mathbb{R})$ und $\hat{\mathbb{R}}^k = \mathtt{Mat}(1,k,\mathbb{R})$.

Wir verallgemeinern die Paarung $\bullet$ zu einem Produkt

$$\bullet : \mathtt{Mat}(n,m,\mathbb{R}) \times \mathtt{Mat}(m,k,\mathbb{R}) \to \mathtt{Mat}(n,k,\mathbb{R}) ,$$

Definition 1.5. *Sei*

$$A = \begin{pmatrix} \hat{a}_1 \\ \vdots \\ \hat{a}_n \end{pmatrix} \in \mathtt{Mat}(n,m,\mathbb{R}) \ , \quad a_i \in \hat{\mathbb{R}}^m$$

und

$$B = (b_1, \ldots, b_k) \in \mathtt{Mat}(m,k,\mathbb{R}) \ , \quad b_i \in \mathbb{R}^m \ .$$

Dann definieren wir

$$A \bullet B := \begin{pmatrix} \hat{a}_1 \bullet b_1 & \ldots & \hat{a}_1 \bullet b_k \\ \vdots & & \vdots \\ \hat{a}_n \bullet b_1 & \ldots & \hat{a}_n \bullet b_k \end{pmatrix} \ .$$

Es gilt also

$$(A \bullet B)_{i,j} = \sum_{l=1}^{m} a_{i,l} b_{l,j} \ .$$

Hier ist ein Beispiel:

$$\begin{pmatrix} 1 & 2 & 3 \\ 0 & 2 & 1 \end{pmatrix} \bullet \begin{pmatrix} 4 & -1 \\ 1 & 0 \\ 2 & 2 \end{pmatrix} = \begin{pmatrix} 1*4+2*1+3*2 & 1*-1+2*0+3*2 \\ 0*4+2*1+1*2 & 0*-1+2*0+1*2 \end{pmatrix} = \begin{pmatrix} 12 & 5 \\ 4 & 2 \end{pmatrix} .$$

Beachte, daß $A \bullet B$ nur dann definiert ist, wenn die Anzahl der Spalten von A mit der Anzahl der Zeilen von B übereinstimmt.

Das Gleichungssystem (3) kann nun wie folgt geschrieben werden. Wir setzen

$$M := \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} a_1 & a_2 \\ 1 & 1 \end{pmatrix} \ , \quad B := \begin{pmatrix} e \\ u \end{pmatrix} \ .$$

Dann lautet die Bestimmungsgleichung

$$M \bullet z = B \ .$$

1.2.4 Assoziativität und nicht-Kommutativität der Matrizenmultiplikation

Seien $n,m,k,l \in \mathbb{N}$ und $A \in \mathtt{Mat}(n,m,\mathbb{R})$, $B \in \mathtt{Mat}(m,k,\mathbb{R})$, und $C \in \mathtt{Mat}(k,l,\mathbb{R})$.

Lemma 1.6. *Es gilt* $(A \bullet B) \bullet C = A \bullet (B \bullet C)$.

Proof. Es gilt

$$(A \bullet B)_{i,y} = \sum_{x=1}^{m} A_{i,x} B_{x,y}$$

und damit

$$((A \bullet B) \bullet C)_{i,j} = \sum_{y=1}^{k} \sum_{x=1}^{m} A_{i,x} B_{x,y} C_{y,j} \ .$$

Auf der anderen Seite gilt

$$(B \bullet C)_{x,j} = \sum_{y=1}^{k} B_{x,y} C_{y,j}$$

und damit

$$(A \bullet (B \bullet C))_{i,j} = \sum_{x=1}^{m} \sum_{y=1}^{k} A_{i,x} B_{x,y} C_{y,j} \ .$$

Wir schließen das Lemma aus

$$\sum_{y=1}^{k} \sum_{x=1}^{m} A_{i,x} B_{x,y} C_{y,j} = \sum_{x=1}^{m} \sum_{y=1}^{k} A_{i,x} B_{x,y} C_{y,j} \ .$$

□

Beachte, daß die Matrixmultiplikation nicht kommutativ ist. Wenn $A, B \in \mathtt{Mat}(n, n, \mathbb{R})$, dann ist zwar $A \bullet B$ und $B \bullet A$ definiert, aber diese beiden Produkte sind im allgemeinen nicht gleich. Hier ist ein einfaches Beispiel.

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

1.2.5 Die Addition von Matrizen und das Distributivgesetz

Der Vollständigkeit halber führen wir gleich noch zwei weitere Operationen mit Matrizen ein.

Definition 1.7. *Wir definieren die Addition* $+ : \mathtt{Mat}(n, m, \mathbb{R}) \times \mathtt{Mat}(n, m, \mathbb{R}) \to \mathtt{Mat}(n, m, \mathbb{R})$ *durch*

$$(A + B)_{i,j} := A_{i,j} + B_{i,j} \ .$$

Lemma 1.8. *Es gilt* $A+(B+C)=(A+B)+C$ *und (wenn die Multiplikation definiert ist)*

$$A\bullet(B+C)=A\bullet B+A\bullet C\ ,\quad (B+C)\bullet D=B\bullet D+C\bullet D\ .$$

Proof. Nachrechnen! □

Sei $\lambda\in\mathbb{R}$ und $A\in \mathtt{Mat}(n,k,\mathbb{R})$.

Definition 1.9. *Wir definieren die skalare Multiplikation* $\lambda A\in\mathtt{Mat}(n,k,\mathbb{R})$ *durch* $(\lambda A)_{i,j}:=\lambda A_{i,j}$.

Wir überlassen es dem Leser, die grundlegenden Eigenschaften dieser Operation zu finden.

1.2.6 Lineare Gleichungen mit drei Unbekannten

Eine quadratische Funktion $f:\mathbb{R}\to\mathbb{R}$ ist eine Funktion der Form $f(x)=a_2x^2+a_1x+a_0$. Analog kann man Funktionen höheren Grades definieren. Wir stellen uns das Problem, eine quadratische Funktion so zu finden, daß sie in drei gegebenen Punkten x_1,x_2,x_3 vorgegebene Werte y_1,y_2,y_3 annimmt. Dies gibt drei Bedingungen für drei Unbekannte a_2,a_1,a_0, die wir dadurch bestimmen wollen.

Wir müssen die Gleichungen

$$a_2x_i^2+a_1x_i+a_0=y_i\ ,\quad i=1,2,3$$

erfüllen. Dies liefert ein lineares Gleichungssystem, welches wir kurz in der Form

$$\begin{pmatrix}1&x_1&x_1^2\\1&x_2&x_2^2\\1&x_3&x_3^2\end{pmatrix}\bullet\begin{pmatrix}a_0\\a_1\\a_2\end{pmatrix}=\begin{pmatrix}y_1\\y_2\\y_3\end{pmatrix}$$

schreiben können.

Es gibt ein systematisches Verfahren, mit welchem man ein derartiges System lösen kann (Gaussverfahren).

Zur Illustration führen wir eine vereinfachte Variante des Verfahrens durch. Eine allgemeine Beschreibung folgt später. Die Idee besteht darin, das System unter Erhaltung

der Lösungsmenge so umzuformen, daß eine Lösung einfach möglich wird. Wir ersetzen zuerst die zweite und dritte Zeile des Systems jeweils durch die Differenz mit der ersten.

$$\begin{pmatrix} 1 & x_1 & x_1^2 \\ 0 & x_2-x_1 & x_2^2-x_1^2 \\ 0 & x_3-x_1 & x_3^2-x_1^2 \end{pmatrix} \bullet \begin{pmatrix} a_0 \\ a_1 \\ a_2 \end{pmatrix} = \begin{pmatrix} y_1 \\ y_2-y_1 \\ y_3-y_1 \end{pmatrix} .$$

Wir setzen jetzt voraus, daß alle x_i paarweise verschieden sind. In unserer Anwendung ist das sinnvoll, da wir ja jeweils nur einen Wert der Funktion in einem Punkt vorgeben können. Wir ersetzen die dritte Zeile durch ihre Differenz mit dem $\frac{x_3-x_1}{x_2-x_1}$-fachen der zweiten.

$$\begin{pmatrix} 1 & x_1 & x_1^2 \\ 0 & x_2-x_1 & x_2^2-x_1^2 \\ 0 & 0 & (x_3-x_1)(x_3-x_2) \end{pmatrix} \bullet \begin{pmatrix} a_0 \\ a_1 \\ a_2 \end{pmatrix} = \begin{pmatrix} y_1 \\ y_2-y_1 \\ y_3-y_1-\frac{x_3-x_1}{x_2-x_1}(y_2-y_1) \end{pmatrix} .$$

Dieses System kann nun leicht durch Rückwärtseinsetzen gelöst werden. Es gilt

$$a_2 = \frac{y_3-y_1-\frac{x_3-x_1}{x_2-x_1}(y_2-y_1)}{(x_3-x_1)(x_3-x_2)} = \frac{y_3(x_2-x_1)+y_1(x_3-x_2)+y_2(x_1-x_3)}{(x_2-x_1)(x_3-x_1)(x_3-x_2)} .$$

Daraus erhält man

$$a_1 = \frac{y_2-y_1-(x_2^2-x_1^2)a_2}{x_2-x_1}$$

und

$$a_0 = y_1-x_1a_1-x_1^2a_2 .$$

Hier ist ein Zahlenbeispiel zur Probe. Sei etwa $x_1=-1$, $x_2=0$ und $x_3=1$ und $y_1=2$, $y_2=1$ und $y_3=2$. Der Graph Lösung sollte die um 1 nach oben verschobene Standardparabel sein. In der Tat erhalten wir $a_2=\frac{2+2-2}{2}=1$, $a_1=\frac{1-2+1}{1}=0$, und $a_0=2+0-1=1$, also $f(x)=x^2+1$.

1.2.7 Widersprüchliche, zu wenige oder zu viele Gleichungen

Es kann durchaus vorkommen, daß ein lineares Gleichungssystem keine Lösung hat. Wenn wir im obigen Beispiel etwa $x_2=x_3$, aber $y_2\neq y_3$ wählen, dann kann es offensichtlich keine Lösung geben.

Wenn aber auch $y_2 = y_3$ gilt, dann kann man die dritte Gleichung weglassen. Wir haben dann das System

$$\begin{pmatrix} 1 & x_1 & x_1^2 \\ 1 & x_2 & x_2^2 \end{pmatrix} \bullet \begin{pmatrix} a_0 \\ a_1 \\ a_2 \end{pmatrix} = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$$

Wir führen den ersten Schritt analog durch und erhalten

$$\begin{pmatrix} 1 & x_1 & x_1^2 \\ 0 & x_2 - x_1 & x_2^2 - x_1^2 \end{pmatrix} \bullet \begin{pmatrix} a_0 \\ a_1 \\ a_2 \end{pmatrix} = \begin{pmatrix} y_1 \\ y_2 - y_1 \end{pmatrix} .$$

Wir können nun a_2 beliebig wählen. Dann ergibt sich

$$a_1 = \frac{y_2 - y_1 - (x_2^2 - x_1^2)a_2}{x_2 - x_1}$$

und

$$a_0 = y_1 - x_1 a_1 - x_1^2 a_2 .$$

In diesem Fall haben wir eine Familie von Lösungen, welche von einem Parameter abhängt.

Es wäre auch denkbar, den Wert y_4 in einem vierten Punkt x_4 vorzuschreiben. Da die bisherigen Daten die Funktion aber schon eindeutig festlegen, hat das System

$$\begin{pmatrix} 1 & x_1 & x_1^2 \\ 1 & x_2 & x_2^2 \\ 1 & x_3 & x_3^2 \\ 1 & x_4 & x_4^2 \end{pmatrix} \bullet \begin{pmatrix} a_0 \\ a_1 \\ a_2 \end{pmatrix} = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{pmatrix}$$

in der Regel keine Lösung (außer die durch die ersten drei Punkte bestimmte Funktion hat zufällig den Wert y_4 im Punkt x_4.)

Mit Hilfe der Theorie der Vektorräume werden wir eine allgemeine Theorie des Lösungsverhaltens linearer Gleichungssysteme entwickeln.

1.2.8 Stochastische Matrizen

Wir legen n Punkte äquidistant auf einen Kreis. Wir stellen uns einen Frosch vor, der auf diesen Punkten lebt und in jedem Zeitschritt $t \in \mathbb{N}$ entweder sitzen bleibt oder zu einem

benachbarten Punkt springt. Wir wissen im einzelnen nicht genau, wie sich der Frosch entscheidet. Nur, daß er mit Wahrscheinlichkeit $1/2$ schläft, und mit Wahrscheinlichkeit $1/2$ springt, wobei beide Richtungen gleichwahrscheinlich sind. Der Zustand (Ort) zur Zeit $t \in \mathbb{N}$ des Frosches muß also durch ein stochastisches Modell beschrieben werden, nämlich durch die Angabe der Wahrscheinlichkeit $P(i,t)$, daß sich der Frosch zur Zeit $t \in \mathbb{N}_0$ am i-ten Punkt befindet. Der Zustand ist also durch $P(t) = (P(t,1),\ldots,P(t,n))^t \in \mathbb{R}^n$ mit den Nebenbedingungen $P(t,i) \in [0,1]$ ($P(i,t)$ ist eine Wahrscheinlichkeit) und $\sum_{i=1}^n P(t,i) = 1$ (es ist sicher, daß sich der Frosch irgendwo befindet) gegeben. Wir können die zweite Bedingung auch kurz durch $\hat{b} \bullet P(t) = 1$ mit $\hat{b} := (1,\ldots,1) \in \hat{\mathbb{R}}^n$ schreiben. Wenn $P(t)$ bekannt ist, dann sind wir in der Lage, $P(t+1)$ zu berechnen. In der Tat, die Wahrscheinlichkeit, zur Zeit t in i und zur Zeit $t+1$ in j zu sein, ist durch $A_{j,i}P(t,i)$ gegeben, wobei $A_{j,i}$ die Wahrscheinlichkeit eines Sprunges von i nach j ist. Folglich ist

$$P(t+1,j) = \sum_{i=1}^{n} A_{j,i}P(t,i) \ .$$

In anderen Worten, wenn $A \in \mathtt{Mat}(n,n,\mathbb{R})$ die Matrix mit den Einträgen $A_{i,j}$ ist, dann ist

$$P(t+1) = A \bullet P(t) \ .$$

In unserem Beispiel ist

$$A = \begin{pmatrix} \frac{1}{2} & \frac{1}{4} & 0 & 0 & \ldots & 0 & 0 & \frac{1}{4} \\ \frac{1}{4} & \frac{1}{2} & \frac{1}{4} & 0 & \ldots & 0 & 0 & 0 \\ 0 & \frac{1}{4} & \frac{1}{2} & \frac{1}{4} & \ldots & 0 & 0 & 0 \\ \vdots & & & & & & \vdots & \\ 0 & 0 & 0 & 0 & \ldots & \frac{1}{4} & \frac{1}{2} & \frac{1}{4} \\ \frac{1}{4} & 0 & 0 & 0 & \ldots & 0 & \frac{1}{4} & \frac{1}{2} \end{pmatrix} \ .$$

Diese Matrix ist eine sogenannte stochastische Matrix, da alle Einträge nicht negativ und die Identität $\hat{b} \bullet A = \hat{b}$ gilt. In der Tat hat das zur Folge, daß $A \bullet P(t)$ wieder in unserem Zustandsraum liegt, da $(A \bullet P(t))_i \geq 0$ und wegen 1.6

$$\hat{b} \bullet (A \bullet P(t)) = (\hat{b} \bullet A) \bullet P(t) = \hat{b} \bullet P(t) = 1$$

gilt. Gibt man den Anfangszustand $P(0)$ vor, dann kann man $P(t)$ durch

$$P(t) = A^t \bullet P(0)$$

explizit ausrechnen. In diesem Fall gilt für $t \to \infty$, daß alle Einträge von A^t gegen $1/n$ konvergieren. Dies bedeutet, daß $P(t)$ für große t gegen eine Gleichverteilung konvergiert unabhängig vom Startwert.

Hier sind ist ein Beispiel:

$$A := \begin{pmatrix} \frac{1}{2} & \frac{1}{4} & 0 & 0 & \frac{1}{4} \\ \frac{1}{4} & \frac{1}{2} & \frac{1}{4} & 0 & 0 \\ 0 & \frac{1}{4} & \frac{1}{2} & \frac{1}{4} & 0 \\ 0 & 0 & \frac{1}{4} & \frac{1}{2} & \frac{1}{4} \\ \frac{1}{4} & 0 & 0 & \frac{1}{4} & \frac{1}{2} \end{pmatrix}$$

$$A^2 = \begin{pmatrix} \frac{3}{8} & \frac{1}{4} & \frac{1}{16} & \frac{1}{16} & \frac{1}{4} \\ \frac{1}{4} & \frac{3}{8} & \frac{1}{4} & \frac{1}{16} & \frac{1}{16} \\ \frac{1}{16} & \frac{1}{4} & \frac{3}{8} & \frac{1}{4} & \frac{1}{16} \\ \frac{1}{16} & \frac{1}{16} & \frac{1}{4} & \frac{3}{8} & \frac{1}{4} \\ \frac{1}{4} & \frac{1}{16} & \frac{1}{16} & \frac{1}{4} & \frac{3}{8} \end{pmatrix}$$

$$A^3 = \begin{pmatrix} \frac{5}{16} & \frac{15}{64} & \frac{7}{64} & \frac{7}{64} & \frac{15}{64} \\ \frac{15}{64} & \frac{5}{16} & \frac{15}{64} & \frac{7}{64} & \frac{7}{64} \\ \frac{7}{64} & \frac{15}{64} & \frac{5}{16} & \frac{15}{64} & \frac{7}{64} \\ \frac{7}{64} & \frac{7}{64} & \frac{15}{64} & \frac{5}{16} & \frac{15}{64} \\ \frac{15}{64} & \frac{7}{64} & \frac{7}{64} & \frac{15}{64} & \frac{5}{16} \end{pmatrix}$$

$$A^6 = \begin{pmatrix} \frac{237}{1024} & \frac{859}{4096} & \frac{715}{4096} & \frac{715}{4096} & \frac{859}{4096} \\ \frac{859}{4096} & \frac{237}{1024} & \frac{859}{4096} & \frac{715}{4096} & \frac{715}{4096} \\ \frac{715}{4096} & \frac{859}{4096} & \frac{237}{1024} & \frac{859}{4096} & \frac{715}{4096} \\ \frac{715}{4096} & \frac{715}{4096} & \frac{859}{4096} & \frac{237}{1024} & \frac{859}{4096} \\ \frac{859}{4096} & \frac{715}{4096} & \frac{715}{4096} & \frac{859}{4096} & \frac{237}{1024} \end{pmatrix}$$

1.2.9 Quadratische Funktionen und Matrizen

Sei $\mathbb{N}_0 := \mathbb{N} \cup \{0\}$. Sei $\mathbb{N}_0^n$ der Raum geordneten der n-Tupel in $\mathbb{N}_0$. Ein Element $\alpha \in \mathbb{N}_0^n$ kann also in der Form $\alpha = (\alpha_1, \ldots, \alpha_n)$ geschrieben werden. Wir definieren

$$|\alpha| := \sum_{i=1}^{n} \alpha_i \, .$$

Für $\alpha \in \mathbb{N}_0^n$ betrachten wir die Funktion

$$x^\alpha : \mathbb{R}^n \to \mathbb{R} \, , x \mapsto x_1^{\alpha_1} \ldots x_n^{\alpha_n} \, .$$

Definition 1.10. *Eine polynomiale Funktion $f : \mathbb{R}^n \to \mathbb{R}$ vom Grade $\leq k$ ist eine Funktion der Form*

$$f(x) = \sum_{\alpha \in \mathbb{N}_0^n, |\alpha| \leq k} a_\alpha x^\alpha$$

mit Koeffizienten $a_\alpha \in \mathbb{R}$.

Dies ist die Verallgemeinerung von Polynomen auf mehrere Veränderliche. Ein Beispiel einer Funktion vom Grad ≤ 4 ist die Funktion $f : \mathbb{R}^3 \to \mathbb{R}$, welche durch

$$f(x_1, x_2, x_3) = 3x_1^2 + x_2 x_3 - 4x_3^4$$

gegeben ist.

Funktionen vom Grad ≤ 0 sind konstant. Funktionen vom Grad ≤ 1 können in der Form

$$f(x) = b \bullet x + c$$

mit $b \in \hat{\mathbb{R}}^n$ und $c \in \mathbb{R}$ geschrieben werden. Dabei ist $b_i = a_{0,\ldots,1,\ldots,0}$ (die 1 steht an der i-ten Stelle) und $c = a_{0,\ldots,0}$.

Der Zusammenhang von quadratischen Funktionen und Matrizen ist durch die Darstellung

$$f(x) = x^t \bullet A \bullet x + b \bullet x + c$$

gegeben, wobei

$$A := \begin{pmatrix} a_{1,1} & \frac{1}{2}a_{1,2} & \ldots & \frac{1}{2}a_{1,n} \\ \frac{1}{2}a_{2,1} & a_{2,2} & \ldots & \frac{1}{2}a_{2,n} \\ \vdots & & & \vdots \\ \frac{1}{2}a_{n,1} & \ldots & \frac{1}{2}a_{n,n1} & a_{n,n} \end{pmatrix} , \quad b = (a_1, \ldots, a_n) .$$

Es gilt hierbei für $i \neq j$

$$a_{i,j} = a_{0,\ldots,1,\ldots,1,\ldots,0} ,$$

wobei die 1 and der i-ten und j-ten Stelle steht. Weiter gilt $a_{i,i} = a_{0,\ldots,2,\ldots,0}$, wobei die 2 an der i-ten Stelle steht. Beachte, den Faktor $1/2$ vor den off-Diagonaleinträgen. Die Matrix A ist symmetrisch, d.h. sie erfüllt $A_{i,j} = A_{j,i}$.

Hier ist eine Anwendung aus der Physik. Wir wollen die kinetische Energie eines um seinen Massenschwerpunkt rotierenden Körpers ausdrücken. Wir beschreiben die Rotation durch ein Tupel (Vektor der Winkelgeschwindigkeiten) $x = (x_1, x_2, x_3)^t$, welches sowohl die Rotationsachse (die Gerade durch x) als auch die Geschwindigkeit bestimmt (die Länge von x). Der Wert von x_i beschreibt den Anteil der Rotation in der Ebene $\{(y_1, y_2, y_3)^t \in \mathbb{R}^3 | y_i = 0\}$. Der genaue Wert ist die Geschwindigkeit im Einheitsabstand von der Achse (in einem orientierten Sinne genommen).

Der Körper bestehe aus durch masselose Streben zusammengehaltene Massepunkte m_α in den Orten $a(\alpha) \in \mathbb{R}^3$. Wir bilden die folgende Matrix

$$M_{i,j} = \sum_\alpha m_\alpha \left[\|a(\alpha)\|^2 \delta_{i,j} - a(\alpha)_i a(\alpha)_j \right] ,$$

wobei das Kroneckersymbol die Werte $\delta_{i,j} := 0$ für $i \neq j$ und $\delta_{i,i} := 1$ hat und $\|a(\alpha)\|^2 := \sum_{i=1}^3 a(\alpha)_i^2$ ist. Diese Matrix beschreibt das Trägheitsverhalten des Körpers bezüglich Rotationen. So ist die kinetische Energie im Rotationszustand $x \in \mathbb{R}^3$ durch

$$E = \frac{1}{2} x^t \bullet M \bullet x .$$

gegeben.

Besteht der Köper etwa aus einem Punkt der Masse m in $(a,0,0)$ und rotiert in der Ebene $x_3 = 0$ mit der Geschwindigkeit $x = (0,0,u)$, dann ist seine kinetische Energie $E = \frac{mu^2a^2}{2}$ nach der obigen Formel. In der Tat ist die absolute Geschwindigkeit des Punktes durch $v = ua$ gegeben, und die Formel reduziert sich auf die bekannte $E = \frac{mv^2}{2}$.

Wenn die Masse 1 im Punkt $(1,1,1)^t$ sitzt, dann gilt

$$M = \begin{pmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{pmatrix} .$$

1.2.10 Widerstandsnetzwerke

Wir betrachten ein Netz aus (elektrischen) Widerständen. Es gebe n Knoten. Der Widerstand zwischen dem i-ten und j-ten Knoten sei $R_{i,j}$. Wir verlangen $R_{i,j} = R_{j,i}$. Seien Spannungen U_i (gegenüber einem Massepotential) in den Knoten vorgegeben. Dann ergibt sich der Strom vom i-ten zum j-ten Knoten durch

$$I_{i,j} = \frac{U_j - U_i}{R_{i,j}} .$$

Der Nettoabfluß im i-ten Knoten ist also durch

$$I_i^{netto} = \sum_j I_{i,j} = \sum_j \frac{U_j - U_i}{R_{i,j}}$$

gegeben. Eine typische Aufgabenstellung besteht jetzt darin, die Spannungen bei vorgegebener Nettostromabflüssen zu bestimmen. Es ergibt sich ein lineares Gleichungssystem

$$B \bullet U = I^{netto}$$

mit

$$B_{i,j} = \frac{1}{R_{i,j}} - \delta_{i,j} \sum_j \frac{1}{R_{i,j}} .$$

Hierbei ist $\delta_{i,j} = 0$ für $i \neq j$ und $\delta_{i,i} = 1$. Ist i ein freier Knoten, so gilt nach dem Erhaltungssatz für Ladungen $I_i^{netto} = 0$. Über dieses System können wir apriori schon einiges sagen. Wenn es eine Lösung U hat, so auch $U + (c,\ldots,c)^t$ für jedes $c \in \mathbb{R}$. Damit überhaupt eine Lösung existiert, muß $\sum_i I_i^{netto} = 0$ gelten. In der Tat kann man zeigen, daß

diese Bedingung hinreichend für die Existenz der Lösung ist, welche dann eindeutig bis auf die Verschiebungen bestimmt ist (dabei sind alle $R_{i,j} \neq \infty$). Wenn einige $R_{i,j} = \infty$ sind (die Knoten sind also nicht verbunden, und wir setzen $\frac{1}{R_{i,j}} = 0$), dann kann das Netzwerk in unabhängige Teilnetzwerke zerfallen, für welche dann entsprechende Existenz- und Eindeutigkeitsaussagen gelten.

Seien alle Widerstände $R_{i,j} = 1$. Dann hat die Matrix B für 4 Knoten die Form

$$\begin{pmatrix} -3 & 1 & 1 & 1 \\ 1 & -3 & 1 & 1 \\ 1 & 1 & -3 & 1 \\ 1 & 1 & 1 & -3 \end{pmatrix}$$

1.3 Der Gauss-Algorithmus

1.3.1 Invertierbare Matrizen

Definition 1.11. *Die Matrix $E_n \in \mathtt{Mat}(n,n,\mathbb{R})$ mit $(E_n)_{i,j} = 0$ falls $i \neq j$ und $(E_n)_{i,i} = 1$ heißt Einheitsmatrix.*

Zur Anschauung:

$$E_n = \begin{pmatrix} 1 & 0 & \dots & \dots & 0 \\ 0 & 1 & 0 & \dots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & 1 & 0 \\ 0 & \dots & \dots & 0 & 1 \end{pmatrix} .$$

Es gilt für $A \in \mathtt{Mat}(n,k,\mathbb{R})$ und $B \in \mathtt{Mat}(m,n,\mathbb{R})$, daß

$$E_n \bullet A = A \, , \quad B \bullet E_n = B \, .$$

Definition 1.12. *Eine Matrix $A \in \mathtt{Mat}(n,n,\mathbb{R})$ heißt invertierbar, wenn es Matrizen $B,C \in \mathtt{Mat}(n,n,\mathbb{R})$ gibt mit der Eigenschaft*

$$A \bullet C = E_n = B \bullet A \, .$$

Lemma 1.13. *Sei A invertierbar und*

$$A \bullet C = E_n = B \bullet A$$

wie in 1.12. Dann gilt $B = C$. Die Matrix B ist durch A eindeutig bestimmt.

Proof. Wir multiplizieren die Gleichung $E_n = A \bullet C$ von links mit B. Dann folgt $B \bullet E_n = B \bullet (A \bullet C)$, also $B = (B \bullet A) \bullet C = E_n \bullet C = C$. Wenn weiter $B' \in \mathtt{Mat}(n,n,\mathbb{R})$ die Gleichung $E_n = B' \bullet A$ erfüllt, so gilt mit dem gleichen Argument $B' = C$, also $B' = C$. □
Wir schreiben oft $A^{-1} := B$ und nennen diese Matrix die inverse Matrix zu A. Beachte, daß A^{-1} auch invertierbar ist und $A = (A^{-1})^{-1}$ gilt.

In der Definition des Begriffs "invertierbar" verlangen wir die Existenz von Links- und Rechtsinversen gleichzeitig. In der Tat ist es ausreichend, nur die Existenz eines dieser beiden zu fordern (siehe **??**). Wir werden dies später aus der Dimensionsformel schließen. Wenn etwa ein Linksinverses existiert, so ist die durch A dargestellte lineare Abbildung $A \bullet \cdots : \mathbb{R}^n \to \mathbb{R}^n$ injektiv. Daraus folgt schon die Surjektivität, und damit die Existenz einer inversen Abbildung, welche eben durch die inverse Matrix dargestellt wird. Genaueres dazu später.

Lemma 1.14. *Wenn $A, B \in \mathtt{Mat}(n,n,\mathbb{R})$ invertierbar sind, dann ist auch $A \bullet B$ invertierbar, und es gilt*

$$(A \bullet B)^{-1} = B^{-1} \bullet A^{-1} \ .$$

Proof. Nachrechnen. □
Beachte die Vertauschung der Reihenfolge!

Hier ist eine einfache Formel für das Inverse einer 2×2-Matrix. Sei

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \ .$$

Dann gilt (wie man leicht durch Nachrechnen einsieht)

$$A^{-1} = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix} \ .$$

Wir müssen hier natürlich voraussetzen, daß die sogenannte Determinante (siehe **??**)

$$\det(A) := ad - bc$$

von A nicht verschwindet. Diese Formel ist ein Spezialfall der Adjunktenformel **??**, welche auch für größere Matrizen angewendet werden kann.

1.3.2 Beispiele für invertierbare Matrizen

Sei $(i,j) \in \overline{n} \times \overline{n}$. Wir betrachten die Matrix $P_n(i,j) \in \mathtt{Mat}(n,n,\mathbb{Z})$, welche durch

$$P_n(i,j)_{x,y} := \begin{cases} 1 & x = y, x \neq i, x \neq j \\ 1 & x \neq y, x = j, y = i \\ 1 & x \neq y, x = i, y = j \\ 0 & \text{sonst} \end{cases}$$

gegeben wird. Ist $A \in \mathtt{Mat}(n,k,\mathbb{R})$, so erhält man $P_n(i,j) \bullet A$ aus A durch Vertauschen der i-ten und j-ten Zeilen. Ist $B \in \mathtt{Mat}(m,n,\mathbb{R})$, dann ergibt sich $B \bullet P_n(i,j)$ durch Vertauschen der i-ten und j-ten Spalten. Dies rechnet man direkt nach. Die Matrix $P_n(i,j)$ ist invertierbar. In der Tat gilt

$$P_n(i,j) \bullet P_n(i,j) = E_n \ .$$

Sei nun $i \neq j$. Für $\lambda \in \mathbb{R}$ betrachten wir die Matrix $E_n(i,j,\lambda)$, welche durch

$$E_n(i,j,\lambda)_{x,y} = \begin{cases} 1 & x = y \\ \lambda & x = i, y = j \\ 0 & \text{sonst} \end{cases}$$

gegeben ist. Ist $A \in \mathtt{Mat}(n,k,\mathbb{R})$, so erhält man $E_n(i,j,\lambda) \bullet A$ aus A, indem man zur i-ten Zeile das λ-fache der j-ten Zeile addiert. Analog, wenn $B \in \mathtt{Mat}(m,n,\mathbb{R})$ ist, dann ergibt sich $B \bullet E_n(i,j,\lambda)$, indem man zur j-ten Spalte das λ-fache der i-ten Spalte addiert. Die Matrix $E_n(i,j,\lambda)$ ist invertierbar. In der Tat gilt

$$E_n(i,j,\lambda) \bullet E_n(i,j,-\lambda) = E_n \ .$$

Zuletzt betrachten wir für $\mathbb{R} \ni \lambda \neq 0$ die Matrix $D_n(i,\lambda) \in \mathtt{Mat}(n,n,\mathbb{R}$, welche durch

$$D(i,\lambda)_{x,y} = \begin{cases} 1 & x = y, x \neq i \\ \lambda & x = i = j \\ 0 & \text{sonst} \end{cases}$$

Ist $A \in \mathtt{Mat}(n,k,\mathbb{R})$, so erhält man $D_n(i,\lambda) \bullet A$ aus A, indem man die i-te Zeile mit λ multipliziert. Analog, wenn $B \in \mathtt{Mat}(m,n,\mathbb{R})$ ist, dann ergibt sich $B \bullet D_n(i,\lambda)$ aus B durch Multiplikation der i-ten Spalte mit λ. Auch diese Matrix ist invertierbar, denn es gilt $D_n(i,\lambda) \bullet D_n(i,\lambda^{-1}) = E_n$.

1.3.3 Die Lösungsmenge eines linearen Gleichungssystems

Seien $k, n \in \mathbb{N}$, $A \in \mathtt{Mat}(n,k,\mathbb{R})$ und $b \in \mathbb{R}^n = \mathtt{Mat}(n,1,\mathbb{R})$. Diese beiden Daten bestimmen ein System (linearer) Gleichungen für $x \in \mathbb{R}^k \cong \mathtt{Mat}(k,1,\mathbb{R})$ durch

$$A \bullet x = b \ .$$

Mit

$$L(A,b) := \{x \in \mathbb{R}^k | A \bullet x = b\} \subset \mathbb{R}^k$$

bezeichnen wir die Menge der Lösungen des Systems. Die Idee des Gauss-Algorithmus ist es, auf systematische Weise das Gleichungssystem in ein einfacheres mit der im wesentlichen gleichen Lösungsmenge umzuformen.

Lemma 1.15. *Ist $C \in \mathtt{Mat}(n,n,\mathbb{R})$ eine invertierbare Matrix, dann ist*

$$L(C \bullet A, C \bullet b) = L(A,b) \ .$$

Proof. Ist $x \in L(A,b)$, dann gilt $A \bullet x = b$. Durch Multiplizieren mit C erhalten wir $C \bullet (A \bullet x) = C \bullet b$, also $(C \bullet A) \bullet x = C \bullet b$. Dies zeigt $L(A,b) \subset L(C \bullet A, C \bullet b)$.

Ist $x \in L(C \bullet A, C \bullet b)$, dann ist $x \in L(C^{-1} \bullet C \bullet A, C^{-1} \bullet C \bullet b) = L(A,b)$. Es folgt $L(C \bullet A, C \bullet b) \subset L(A,b)$. □

Sei nun $C \in \mathtt{Mat}(k,k,\mathbb{R})$ invertierbar. Dann haben wir eine Bijektion

$$C \bullet \cdots : \mathbb{R}^k \to \mathbb{R}^k \ .$$

In der Tat wird die inverse Abbildung durch C^{-1} gegeben. Wir schreiben $C \bullet L(A,b)$ für die Menge, welche sich durch elementweise Anwendung von $C \bullet \ldots$ auf $L(A,b)$ ergibt.

Lemma 1.16. *Es gilt $C \bullet L(A,b) = L(A \bullet C^{-1}, b)$.*

Proof. Nachrechnen. □

Die Kenntnis von $C \bullet L(A,b)$ ist praktisch gleichwertig mit der Kenntnis von $L(A,b)$, denn man braucht ja nur $C^{-1} \bullet \ldots$ anwenden.

1.3.4 Systeme in Normalform

Sei $A \in \mathtt{Mat}(n,k,\mathbb{R})$.

Definition 1.17. *Wir sagen, daß A in Normalform (vom Rang l) ist, wenn es ein $l \in \{1,\ldots,\mathtt{min}(k,n)\}$ mit*

$$A_{i,j} = \begin{cases} 1 & i = j, i \leq l \\ 0 & \textit{sonst} \end{cases}$$

gibt.

Wenn A in Normalform vom Rang l ist, dann kann man die Lösungsmenge $L(A,b)$ sofort bestimmen. Ist $b_i \neq 0$ für ein $l < i < k$, dann ist $L(A,b) = \emptyset$. Andernfalls ist

$$L(A,b) = \{x \in \mathbb{R}^k | x_i = b_i \text{ für } 1 \leq i \leq l\} \ .$$

Der Gauss-Algorithmus ist ein systematisches Verfahren zur Umformung des Systems in eines in Normalform, wobei die Lösungsmenge (im wesentlichen) erhalten bleibt.

1.3.5 Der Gauss-Algorithmus

Der *Input* des Gaussalgorithmusses ist eine Matrix $A \in \mathtt{Mat}(n,k,\mathbb{R})$ und eine Spalte $b \in \mathbb{R}^n$.

Die *Ausgabe* des Algorithmusses ist eine Matrix $\bar{A} \in \mathtt{Mat}(n,k,\mathbb{R})$, eine Spalte $\bar{b} \in \mathbb{R}^n$ und eine invertierbare Matrix $C \in \mathtt{Mat}(k,k,\mathbb{R})$. Dabei gilt

$$L(A,b) = C \bullet L(\bar{A},\bar{b}) \ .$$

Der Initialisierungsschritt erzeugt

$$A(0) := A \ , \quad b(0) := b \ , C(0) := E_k \ .$$

Der Algorithmus funktioniert iterativ. Die Eingabe des p-ten Schrittes ist $A(p)$, $b(p)$, $C(b)$. Wir nehmen an, daß $A(p)_{i,i} = 1$ für $1 \leq i \leq p$ gilt, daß $A(p)_{i,j} = 0$ für $i \neq j$ und $i \leq p$ oder $j \leq p$, und daß

$$L(A,b) = C(p) \bullet L(A(p),b(p))$$

gilt. Das Ergebnis des Hauptschrittes des Verfahrens sind $A(p+1)$, $b(p+1)$ und $C(p+1)$, welche ebenfalls diese Annahmen erfüllen (wobei natürlich p durch $p+1$ zu ersetzen ist).

Ist $A(p)$ in Normalform, dann ist der Algorithmus beendet. Wir setzen $\overline{A} := A(b)$, $\overline{b} := b(p)$ und $C := C(p)$.

Andernfalls finden wir ein Paar (i, j) mit $p < i$, $p < j$ und $A(p)_{i,j} \neq 0$. Die folgenden Schritte erhalten die Struktur der ersten p Zeilen und Spalten von A. Wir setzen zuerst

$$\begin{aligned} A^{(1)} &:= D_n(p+1, A(p)_{i,j}^{-1}) \bullet P_n(i, p+1) \bullet A(p) \bullet P_k(j, p+1) \\ b^{(1)} &:= D_n(p+1, A(p)_{i,j}^{-1}) \bullet P_n(i, p+1) \bullet b(p) \end{aligned}$$

und

$$C^{(1)} := C(p) \bullet P_k(j, p+1) \ .$$

Nun ist $A^{(1)}_{p+1,p+1} = 1$. Im zweiten Schritt bilden wir

$$\begin{aligned} L &:= E_n(n, p+1, -A^{(1)}_{n,p+1}) \bullet E_n(n-1, p+1, -A^{(1)}_{n-1,p+1}) \bullet \cdots \bullet E_n(p+2, p+1, -A^{(1)}_{p+2,p+1}) \\ A^{(2)} &:= L \bullet A^{(1)} \\ b(p+1) &:= L \bullet b^{(1)} \end{aligned}$$

Diese Umformung produziert die geforderten Nullen in der $p+1$-sten Spalte. Im dritten Schritte produzieren wir die Nullen in der $p+1$-sten Zeile durch

$$\begin{aligned} R &:= E_k(p+1, k, -A^{(2)}_{p+1,k}) \bullet E_k(p+1, k-1, -A^{(2)}_{p+1,k-1}) \bullet \cdots \bullet E_k(p+1, p+2, -A^{(2)}_{p+1,p+2}) \\ A(p+1) &:= A^{(2)} \bullet R \\ C(p+1) &:= C^{(1)} \bullet R \end{aligned}$$

Man prüft nun nach, daß $A(p+1), b(p+1), C(p+1)$ die Annahmen erfüllen.

1.3.6 Ein Zahlenbeispiel

Sei

$$A := \begin{pmatrix} -2 & 1 & 1 \\ 1 & -2 & 1 \\ 1 & 1 & -2 \end{pmatrix} \ , \quad b := \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix} \ .$$

Wir fangen an:

- $p = 0$

$$A(0) = \begin{pmatrix} -2 & 1 & 1 \\ 1 & -2 & 1 \\ 1 & 1 & -2 \end{pmatrix} .$$

$$C(0) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} ,$$

$$b(0) = (1, -2, 1)^t .$$

- Wählen das Paar $(1,1)$. Es ergibt sich

$$A^{(1)} = \begin{pmatrix} 1 & -\frac{1}{2} & -\frac{1}{2} \\ 1 & -2 & 1 \\ 1 & 1 & -2 \end{pmatrix} ,$$

$$C^{(1)} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} ,$$

$$b^{(1)} = \begin{pmatrix} -\frac{1}{2} \\ -2 \\ 1 \end{pmatrix} .$$

- Wir haben

$$L := \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -1 & 0 & 1 \end{pmatrix} \bullet \begin{pmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ -1 & 0 & 1 \end{pmatrix}$$

$$A^{(2)} = \begin{pmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ -1 & 0 & 1 \end{pmatrix} \bullet \begin{pmatrix} 1 & -\frac{1}{2} & -\frac{1}{2} \\ 1 & -2 & 1 \\ 1 & 1 & -2 \end{pmatrix} = \begin{pmatrix} 1 & -\frac{1}{2} & -\frac{1}{2} \\ 0 & -\frac{3}{2} & \frac{3}{2} \\ 0 & \frac{3}{2} & -\frac{3}{2} \end{pmatrix} ,$$

$$b(1) := \begin{pmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ -1 & 0 & 1 \end{pmatrix} \bullet \begin{pmatrix} -\frac{1}{2} \\ -2 \\ 1 \end{pmatrix} = \begin{pmatrix} -\frac{1}{2} \\ -\frac{3}{2} \\ \frac{3}{2} \end{pmatrix}$$

- Schließlich erhalten wir

$$R = \begin{pmatrix} 1 & 0 & \frac{1}{2} \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \bullet \begin{pmatrix} 1 & \frac{1}{2} & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & \frac{1}{2} & \frac{1}{2} \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

$$A(1) = \begin{pmatrix} 1 & -\frac{1}{2} & -\frac{1}{2} \\ 0 & -\frac{3}{2} & \frac{3}{2} \\ 0 & \frac{3}{2} & -\frac{3}{2} \end{pmatrix} \bullet \begin{pmatrix} 1 & \frac{1}{2} & \frac{1}{2} \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -\frac{3}{2} & \frac{3}{2} \\ 0 & \frac{3}{2} & -\frac{3}{2} \end{pmatrix}$$

$$C(1) = \begin{pmatrix} 1 & \frac{1}{2} & \frac{1}{2} \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

- Nun ist $p = 1$. Wir wählen $(i, j) = (2, 2)$. Dann ist

$$A^{(1)} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & -1 \\ 0 & \frac{3}{2} & -\frac{3}{2} \end{pmatrix}$$

$$C^{(1)} = \begin{pmatrix} 1 & \frac{1}{2} & \frac{1}{2} \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

und

$$b^{(1)} = \begin{pmatrix} -\frac{1}{2} \\ 1 \\ \frac{3}{2} \end{pmatrix}$$

- Wir haben nun

$$L = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -\frac{3}{2} & 1 \end{pmatrix}$$

$$A^{(2)} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -\frac{3}{2} & 1 \end{pmatrix} \bullet \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & -1 \\ 0 & \frac{3}{2} & -\frac{3}{2} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{pmatrix}$$

$$b(2) = \begin{pmatrix} -\frac{1}{2} \\ 1 \\ 0 \end{pmatrix}$$

- Schließlich

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

$$A(2) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

$$C(2) = \begin{pmatrix} 1 & \frac{1}{2} & \frac{1}{2} \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \bullet \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & \frac{1}{2} & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$$

Wir sehen also, daß

$$L(A(2), b(2)) = \{(-\frac{1}{2}, 1, x)^t | x \in \mathbb{R}\} \ .$$

Damit ist

$$L(A,b) = C(2)\{(-\frac{1}{2}, -1, x)^t | x \in \mathbb{R}\} = \{(x, 1+x, x)^t | x \in \mathbb{R}\}$$

1.3.7 Bestimmung der inversen Matrix

Der Gauss-Algorithmus liefert auch ein systematisches Verfahren zur Bestimmung von A^{-1}. Sei $A \in \mathtt{Mat}(n,n,\mathbb{R})$. Dann suchen wir eine Matrix $X \in \mathtt{Mat}(n,n,\mathbb{R})$ mit $A \bullet X = E_n$. Sei X_i die i-te Spalte von X und e_i die i-te Spalte von E_n (diese hat lauter Nullen außer der 1 an der i-ten Stelle).

Definition 1.18. *Die Menge $\{e_1, \ldots, e_n\}$ heißt Standardbasis von $\mathbb{R}^n$.*

Wenn wir A, e_i in den Gaussalgorithmus eingeben, dann sei $\overline{A}, \overline{e_i}, C$ die Ausgabe. Wir sehen durch Inspektion ein, daß $\overline{A}$ und C nicht von i abhängen.

Wir nehmen an, daß $\overline{A}$ eine Normalform vom Rang p hat. Wenn $p < n$ ist, und wenn es für alle i mindestens eine Lösung des Systems $A \bullet X_i = e_i$ gibt, so gäbe es mehrere verschiedene Lösungen für X. Dies steht im Widerspruch zur eindeutigen Bestimmtheit der inversen Matrix. Folglich kann dieser Fall nicht eintreten. Wenn $p < n$, dann besitzt A kein Inverses.

Wir nehmen jetzt an, daß $p = n$, also $\overline{A} = E_n$ ist. Dann ist $L(\overline{A}, \overline{e_i}) = \{\overline{e_i}\}$. Folglich gilt

$$A \bullet C \bullet \overline{e_i} = e_i \ .$$

Sei $W := (\overline{e_1}, \ldots, \overline{e_n}) \in \mathtt{Mat}(n, n, \mathbb{R})$ die Matrix mit den Spalten $\overline{e_i}$. Dann gilt offensichtlich

$$A \bullet C \bullet W = E_n \ .$$

Wir sehen also, daß

$$A^{-1} = C \bullet W \ .$$

Hier ist ein Zahlenbeispiel. Sei

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

Wir fügen A mit einer Einheitsmatrix zusammen und erhalten

$$\left(\begin{array}{ccc|ccc} 1 & 2 & 3 & 1 & 0 & 0 \\ 2 & 0 & 1 & 0 & 1 & 0 \\ 0 & 2 & 1 & 0 & 0 & 1 \end{array}\right) \ .$$

Diese Schreibweise ist günstig, um alle rechten Seiten e_i gleichzeitig zu behandeln. Wir führen die Zeilenoperationen wie gehabt aus. Die Spaltenoperationen betreffen nur die ersten drei Spalten.

Die Zeilenoperationen des ersten Schrittes ergeben

$$\left(\begin{array}{ccc|ccc} 1 & 2 & 3 & 1 & 0 & 0 \\ 0 & -4 & -5 & -2 & 1 & 0 \\ 0 & 2 & 1 & 0 & 0 & 1 \end{array}\right) \ .$$

Die Spaltenoperationen des ersten Schrittes liefern

$$\left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & -4 & -5 & -2 & 1 & 0 \\ 0 & 2 & 1 & 0 & 0 & 1 \end{array}\right) \ , \quad C(1) := \begin{pmatrix} 1 & -2 & -3 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \ .$$

Wir normieren.

$$\left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & \frac{5}{4} & \frac{1}{2} & -\frac{1}{4} & 0 \\ 0 & 2 & 1 & 0 & 0 & 1 \end{array}\right) \ .$$

Die Zeilenoperationen des zweiten Schrittes liefern

$$\left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & \frac{5}{4} & \frac{1}{2} & -\frac{1}{4} & 0 \\ 0 & 0 & -\frac{3}{2} & -1 & \frac{1}{2} & 1 \end{array}\right).$$

Die Spaltenoperationen des zweiten Schrittes ergeben

$$\left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & \frac{1}{2} & -\frac{1}{4} & 0 \\ 0 & 0 & -\frac{3}{2} & -1 & \frac{1}{2} & 1 \end{array}\right),$$

$$C(2) = \begin{pmatrix} 1 & -2 & -3 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \bullet \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & -\frac{5}{4} \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & -2 & -\frac{1}{2} \\ 0 & 1 & -\frac{5}{4} \\ 0 & 0 & 1 \end{pmatrix}$$

Die letzte Normierung ergibt

$$\left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & \frac{1}{2} & -\frac{1}{4} & 0 \\ 0 & 0 & 1 & \frac{2}{3} & -\frac{1}{3} & -\frac{2}{3} \end{array}\right).$$

Wir erhalten

$$A^{-1} = \begin{pmatrix} 1 & -2 & -\frac{1}{2} \\ 0 & 1 & -\frac{5}{4} \\ 0 & 0 & 1 \end{pmatrix} \bullet \begin{pmatrix} 1 & 0 & 0 \\ \frac{1}{2} & -\frac{1}{4} & 0 \\ \frac{2}{3} & -\frac{1}{3} & -\frac{2}{3} \end{pmatrix} = \begin{pmatrix} -\frac{1}{3} & \frac{2}{3} & \frac{1}{3} \\ -\frac{1}{3} & \frac{1}{6} & \frac{5}{6} \\ \frac{2}{3} & -\frac{1}{3} & -\frac{2}{3} \end{pmatrix}.$$

Hier ist die Probe.

$$\begin{pmatrix} 1 & 2 & 3 \\ 2 & 0 & 1 \\ 0 & 2 & 1 \end{pmatrix} \bullet \begin{pmatrix} -\frac{1}{3} & \frac{2}{3} & \frac{1}{3} \\ -\frac{1}{3} & \frac{1}{6} & \frac{5}{6} \\ \frac{2}{3} & -\frac{1}{3} & -\frac{2}{3} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

1.3.8 Dreiecksmatrizen

Definition 1.19. *Eine Matrix $A \in \mathtt{Mat}(n,n,\mathbb{R})$ heißt untere (obere) Dreiecksmatrix, wenn für $i > j$ (bzw. $i < j$) die Bedingung $A_{i,j} = 0$ erfüllt ist.*

Hier ist ein Beispiel einer oberen Dreiecksmatrix.

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

Ist A eine Dreiecksmatrix, dann kann das Gleichungssystem $A \bullet x = b$ durch rückwärts Einsetzen gelöst werden. Sei z.B. A eine obere Dreiecksmatrix. Dann ist

$$x_{n-k} = \frac{b_{n-k} - \sum_{i=0}^{k-1} A_{n-k,n-i} x_{n-i}}{A_{n-k,n-k}} .$$

Durch diese Formel können die x_i der Reihe nach beginnend mit $i = n$ und dann im Index absteigend bestimmt werden.

Insbesondere sehen wir sofort, daß unter der Voraussetzung $A_{i,i} \neq 0$ für alle $i = 1, \ldots, n$ in jedem Fall eine Lösung existiert. Es folgt:

Lemma 1.20. *Ist $A \in \mathtt{Mat}(n,n,\mathbb{R})$ eine obere (untere) Dreiecksmatrix, bei der alle Diagonalelemente nicht verschwinden, so ist A invertierbar, wobei A^{-1} ebenfalls eine obere (untere) Dreiecksmatrix ist.*

Proof. Die $n-j$-te Spalte von $C = A^{-1}$ ergibt sich aus

$$C_{n-k,n-j} = \frac{(E_n)_{n-j,n-k} - \sum_{i=0}^{k-1} A_{n-k,n-i} C_{n-i,n-j}}{A_{n-k,n-k}} .$$

Insbesondere ist $C_{n-k,n-j} = 0$ für $k < j$, also für $n-j > n-k$. □

Typische untere (obere) Dreiecksmatrizen sind etwa die $E_n(i,j,\lambda)$ für $i > j$ (bzw. $i < j$) und Diagonalmatrizen. Ein Analyse des Gaussalgorithmus zeigt, daß man eine gegebene Matrix A oft als Produkt von Dreiecksmatrizen schreiben kann. In der Tat, wenn man bei der Durchführung des Algorithmus ohne die Vertauschungsmatrizen $P_n(i,j)$ auskommt, dann gilt für das Endergebnis $\bar{A}$

$$\bar{A} = L \bullet A \bullet R ,$$

wobei L eine untere und R eine obere invertierbare Dreiecksmatrix ist. Da $\bar{A}$ diagonal ist, ist

$$A = (L^{-1} \bullet \bar{A}) \bullet R^{-1}$$

eine derartige Darstellung.

1.4 Geometrische Interpretation linearer Gleichungen

1.4.1 2-dimensionaler Fall

Wir können den $\mathbb{R}^2$ als Modell der (physikalischen) Ebene interpretieren. Dazu führen wir in der physikalischen Ebene ein Koordinatennetz ein, mit dessen Hilfe jeder physikalische Punkt durch ein Paar reeller Zahlen beschrieben werden kann. Wir haben weiter eine Vorstellung von dem Begriff der Geraden, Parallelverschiebung und dem Schnittverhalten mehrerer Geraden.

Diese Dinge haben mathematische Modelle (mit dem gleichen Namen).

Definition 1.21. *Den Raum $\mathbb{R}^2$ bezeichnen wir auch als Ebene.*

Definition 1.22. *Zwei Punkte $x,y \in \mathbb{R}^2$, $x \neq y$ bestimmen die Teilmenge*

$$G(x,y) = \{x + (y-x)t | t \in \mathbb{R}\} \subset \mathbb{R}^2 \ .$$

Eine Gerade in $\mathbb{R}^2$ ist eine Teilmenge dieser Form.

Lemma 1.23. *Sei $G \subset \mathbb{R}^2$ eine Teilmenge. Dann sind folgende Aussagen äquivalent.*

1. *G ist eine Gerade.*
2. *Es gibt $0 \neq \hat{a} \in \hat{\mathbb{R}}^2$ und $b \in \mathbb{R}$ derart, daß $G = L(\hat{a},b)$ ist.*

Proof. Sei $G = G(x,y)$ für $x,y \in \mathbb{R}^2$ mit $x \neq y$. Sei $y - x = (c_1,c_2)^t$. Dann setzen wir $\hat{a} := (c_2,-c_1)$ und $b := \hat{a} \bullet x$. Wir zeigen, daß $G = L(\hat{a},b)$ gilt. In der Tat gilt $\hat{a} \bullet (y-x) = 0$. Wir rechnen mit den Regeln (1)

$$\hat{a} \bullet (x + (y-x)t) = \hat{a} \bullet x + t\hat{a} \bullet (y-x) = b \ .$$

Folglich $G \subset L(\hat{a},b)$. Sei umgekehrt $z \in L(\hat{a},b)$. Wir suchen $t \in \mathbb{R}$ derart, daß $z = x + (y-x)t$. Sei$y_1 - x_1 \neq 0$ (sonst betrachten wir die zweite Koordinate). Dann muß

$$t = \frac{z_1 - x_1}{y_1 - x_1}$$

gelten. Wir betrachten nun die zweite Komponente. Es muß gelten:

$$z_2 = x_2 + (y_2 - x_2)\frac{z_1 - x_1}{y_1 - x_1} \ .$$

Dies ist zu

$$(y_1 - x_1)z_2 - (y_2 - x_2)z_1 = -(y_2 - x_2)x_1 + (y_1 - x_2)x_1$$

äquivalent. Diese Gleichung gilt genau wegen $z \in L(\hat{a}, b)$. Also gilt auch $L(\hat{a}, b) \subset G$.

Es bleibt die Umkehrung zu zeigen. Wir zeigen, daß $L(\hat{a}, b)$ mindestens zwei verschiedene Punkte x, y enthält. Wir nehmen an, daß $\hat{a} = (a_1, a_2)$ und $a_2 \neq 0$ ist. Dann setzen wir $x = (\frac{b}{a_1}, 0)$ und $y = (\frac{b-a_2}{a_1}, 1)$. Es gilt $x, y \in L(\hat{a}, b)$. Dann gilt aber auch $G(x,y) \subset L(\hat{a}, b)$. In der Tat ist

$$\hat{a} \bullet (x + (y - x)t) = (1-t)\hat{a} \bullet x + t\hat{a} \bullet y = (1-t)a_1 \frac{b}{a_1} + t(a_1 \frac{b - a_2}{a_1} + a_2) = b \ .$$

Umgekehrt rechnet man wie oben nach, daß $L(\hat{a}, b) \subset G(x,y)$. □

Ist G eine Gerade, so sind durch $G = L(\hat{a}, b)$ die Einträge $\hat{a}$ und b keineswegs eindeutig bestimmt. So gilt etwa auch $G = L(\mu\hat{a}, \mu b)$ für $0 \neq \mu \in \mathbb{R}$.

Seien G_1, G_2 Geraden. In der physikalischen Welt schneiden sich zwei Geraden entweder in genau einem Punkt, oder sie sind parallel. Im letzteren Fall schneiden sie sich entweder gar nicht, oder sie fallen zusammen. Dasselbe Verhalten zeigen die mathematischen Geraden. Seien $G_i = L(\hat{a}_i, b_i)$, $i = 1, 2$. Dann ist offensichtlich

$$G_1 \cap G_2 = L(A, B)$$

mit

$$A = \begin{pmatrix} \hat{a}_1 \\ \hat{a}_2 \end{pmatrix} \ , \quad B = \begin{pmatrix} b_1 \\ \hat{b}_2 \end{pmatrix} \ .$$

Das Schnittverhalten entspricht also dem Lösungsverhalten des linearen Gleichungssystems

$$A \bullet x = B \ .$$

Wir werden in der Tat sehen, daß dieses System in der Regel genau eine Lösung hat (die Geraden sind nicht parallel). Im Ausnahmefall hat es entweder keine Lösung (die Geraden sind parallel und nicht gleich), oder eine 1-Parameterfamilie von Lösungen (die Geraden sind gleich).

1.4.2 3-dimensionaler Fall

Im physikalischen dreidimensionalen Raum betrachten wir wiederum Geraden, aber auch Ebenen. Durch Wahl eines Koordinatensystems kann man Punkte im Raum durch Tripel reeller Zahlen darstellen. Unser Modell für den (dreidimensionalen) Raum ist $\mathbb{R}^3$.

Definition 1.24. *Der dreidimensionale Raum ist* $\mathbb{R}^3$.

Die Erfahrung, daß zwei verschiedene Punkte eine Gerade bestimmen sollte, legt uns die folgende Definition nahe.

Definition 1.25. *Zwei Punkte* $x,y \in \mathbb{R}^3$, $x \neq y$, *bestimmen eine Gerade*

$$G(x,y) := \{x+(y-x)t | t \in \mathbb{R}\} \subset \mathbb{R}^3 \ .$$

Eine Gerade im $\mathbb{R}^3$ *ist eine Teilmenge dieser Form.*

Der Begriff einer Ebene im $\mathbb{R}^3$ ist etwas komplizierter. Die Definition abstrahiert die Erfahrung, daß eine Ebene durch drei Punkte bestimmt wird.

Definition 1.26. *Sind* $x,y,z \in \mathbb{R}^3$ *Punkte, welche nicht in einer Geraden liegen, dann bestimmen diese Punkte eine Ebene*

$$E(x,y,z) := \{x+t(y-x)+s(z-x) | s,t \in \mathbb{R}\} \ .$$

Eine Ebene im $\mathbb{R}^3$ *ist eine Teilmenge dieser Form.*

Das folgende Lemma wollen wir aus Zeitgründen nicht beweisen und empfehlen es als Übungsaufgabe.

Lemma 1.27. *Sei* $E \subset \mathbb{R}^3$ *eine Teilmenge. Dann sind folgende Aussagen äquivalent.*

1. *E ist eine Ebene.*
2. *Es gibt ein* $0 \neq \hat{a} \in \hat{\mathbb{R}}^3$ *und* $b \in \mathbb{R}$ *derart, daß* $E = L(\hat{a},b)$.

Seien $x,y,z \in \mathbb{R}^k$. Wie prüft man nach, ob x,y,z nicht in einer Geraden liegen (damit etwa $E(x,y,z)$ definiert ist)? Nun, die Punkte x,y bestimmen die Gerade $G(x,y) = \{x+(y-x)t | t \in \mathbb{R}\}$. Es gilt $z \in G(x,y)$ genau dann, wenn es ein $t \in \mathbb{R}$ mit $x+(y-x)t = z$ gibt.

Wir erhalten also ein System $(y-x) \bullet t = (z-x)$ aus drei linearen Gleichungen mit einer Unbekannten. Es gilt $z \in G(x,y)$ genau dann, wenn dieses System eine Lösung besitzt.

Aus der Erfahrung wissen wir, daß sich zwei Ebenen im allgemeinen in einer Gerade schneiden sollten. Falls nicht, dann nennen wir die Ebenen parallel.

Seien $E_i = L(\hat{a}_i, b_i)$, $i = 1,2$. Dann ist $E_1 \cap E_2 = L(A,B)$ mit

$$A = \begin{pmatrix} \hat{a}_1 \\ \hat{a}_2 \end{pmatrix}, \quad B = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}.$$

Das Schnittverhalten von Ebenen entspricht also dem Lösungsverhalten von entsprechenden linearen Gleichungssystemen.

Wir sehen, daß sich Geraden im $\mathbb{R}^3$ als Schnitt zweier Ebenen darstellen lassen, also als Lösungsmenge eines Systems aus zwei linearen Gleichungen. Wir wissen weiter, daß sich zwei Geraden im $\mathbb{R}^3$ im allgemeinen nicht schneiden. In der Tat wäre der Durchschnitt die Lösungsmenge eines Systems aus 4 linearen Gleichungen mit drei Unbekannten, und solche Systeme haben in der Regel keine Lösung. Natürlich gibt es hier jede Menge Ausnahmen von der Regel (wenn etwa beide Geraden in einer Ebene enthalten sind), welche wir im einzelnen nicht diskutieren wollen.

1.4.3 Beliebige Dimension

Die mathematische Abstraktion erlaubt nun, diese Begriffe in noch höheren Dimensionen zu verstehen.

Definition 1.28. *Der k-dimensionale Raum ist* $\mathbb{R}^k$.

Die Teilobjekte sind nun n-dimensionale affine Unterräume, welche für $n = 0$ Punkte, für $n = 1$ Geraden, für $n = 2$ Ebenen und für $n = k-1$ Hyperebenen heißen.

Die folgende Definition bestimmt diese induktiv nach der Dimension.

Definition 1.29. *Eine Folge von Punkten* $x_0, \ldots, x_n \in \mathbb{R}^k$*, welche nicht in einem* $n-1$*-dimensionalen affinen Unterraum liegt, bestimmt einen n-dimensionalen affinen Unterraum*

$$U(x_0, \ldots, x_n) := \{x_0 + \sum_{i=1}^{n} (x_i - x_0) s_i | s \in \mathbb{R}^n\} \subset \mathbb{R}^k .$$

Ein n-dimensionaler Unterraum ist eine Teilmenge dieser Form.

Die Bedingung, daß $x_0, \ldots, x_n$ nicht in einem $n-1$-dimensionalen affinen Unterraum liegen, kann wiederum durch die Lösbarkeit von linearen Gleichungssystemen beschrieben werden. Wir betrachten die Matrix $A \in \mathtt{Mat}(k, n-1, \mathbb{R})$,

$$A := (x_1 - x_0, \ldots, x_{n-1} - x_0) \ .$$

Lemma 1.30. *Wir nehmen an, daß $x_0, \ldots, x_{n-1}$ nicht in einem $n-2$-dimensionalen Unterraum liegen. Dann liegen $x_0, \ldots, x_n$ genau dann in einem $n-1$-dimensionalen affinen Unterraum, wenn $L(A, x_n - x_0) \neq \emptyset$ gilt.*

Proof. (*Nur Notwendigkeit*) Wenn das System eine Lösung hat, dann liegen die Punkte in einem $n-1$-dimensionalen affinen Unterraum. In der Tat, sei $t \in L(A, x_n)$. Dann gilt $x_n = x_0 + \sum_{i=1}^{n-1}(x_i - x_0)t_i$, also $x_n \in U(x_0, \ldots, x_{n-1})$. Ein Beweis der Umkehrung wäre an dieser Stelle zu aufwending. Mit Hilfe der Vektorraumtheorie werden wir die Tatsache später leicht einsehen können. □

Ein Unterraum ohne das Adjektiv “affin” muß den Nullpunkt enthalten (um konform mit der späteren Definition eines Vektorunterraumes zu sein). Mit dem Adjektiv “affin” zeigen wir an, daß wir auch verschobene Unterräume betrachten, welche den Nullpunkt nicht enthalten müssen.

Wir werden später sehen, daß man Unterräume durch lineare Gleichungssysteme beschreiben kann. Dabei entsprechen n-dimensionale Unterräume Systemen mit $k-n$ Gleichungen. Ebenso kann man das Schnittverhalten durch die Lösungstheorie von Gleichungssystemen erkären. Dazu benötigen wir jedoch etwas Theorie der Vektorräume.

2 Algebraische Strukturen : Gruppen, Ringe, Körper

2.1 Gruppen

2.1.1 Die Strukturen einer Gruppe

Definition 2.1. *Eine Menge mit einer Verknüpfung ist ein Paar (M,μ) aus einer Menge M und einer Abbildung*

$$\mu : M \times M \to M \ .$$

Wir haben hiervon schon eine ganze Reihe von Beispielen gesehen.

1. Die Zahlen $\mathbb{N},\mathbb{Z},\mathbb{Q},\mathbb{R}$ haben sogar zwei Verknüpfungen, nämlich die Addition $+$ und die Multiplikation $*$.
2. Die Menge $\mathbb{R}^k$ hat die Addition $+$ als Verknüpfung.
3. Die Menge der Matrizen $\mathtt{Mat}(n,n,R)$ mit $R \in \{\mathbb{N},\mathbb{Z},\mathbb{Q},\mathbb{R}\}$ hat zwei Verknüpfungen, nämlich $\bullet$ und $+$.
4. Sei A eine Menge. Dann hat die Menge der Abbildungen $\mathtt{Hom}(A,A)$ eine Verknüpfung, die Komposition $\circ$.
5. Wir definieren auf $[3] = \{1,2,3\}$ eine Verknüpfung durch die Wertetabelle

	1	2	3
1	3	2	1
2	1	2	1
3	3	3	2

   .

   Also etwa $\mu(1,1) = 3$ und $\mu(3,2) = 3$.

Sei (M,μ) eine Menge mit Verknüpfung und $N \subset M$ eine Teilmenge.

Definition 2.2. *N heißt abgeschlossen unter μ, wenn $\mu(N \times N) \subset N$ gilt (d.h, aus $x,y \in N$ folgt $\mu(x,y) \in N$).*

Wenn $N \subset M$ abgeschlossen unter μ ist, dann kann man μ auf N einschränken und erhält eine Verknüpfung auf N.

1. Sei $GL(n,\mathbb{R}) \subset \mathtt{Mat}(n,n,\mathbb{R})$ die Teilmenge der invertierbaren Matrizen. Dann ist $GL(n,\mathbb{R})$ abgeschlossen unter $\bullet$. $GL(n,\mathbb{R})$ ist jedoch nicht abgeschlossen unter $+$, da etwa $-E_n, E_n \in GL(n,\mathbb{R})$, aber $0 = -E_n + E_n \notin GL(n,\mathbb{R})$ gilt.

2. Sei $S(A) \subset \mathtt{Hom}(A,A)$ die Menge der Bijektionen. Dann ist $S(A)$ abgeschlossen unter $\circ$.

3. Sei $D^{\geq} \subset \mathtt{Mat}(n,n,\mathbb{R})$ die Teilmenge der oberen Dreiecksmatrizen. Dann ist $D^{\geq}$ abgeschlossen unter $+$ und unter $\bullet$.

Definition 2.3. *Eine Verknüpfung auf M heißt assoziativ, falls*

$$\mu(\mu(a,b),c) = \mu(a,\mu(b,c))$$

gilt.

Dies kann man auch durch das Diagramm

$$\begin{array}{ccc} M \times M \times M & \overset{\mu \times \mathtt{id}}{\rightarrow} & M \times M \\ \mathtt{id} \times \mu \downarrow & & \mu \downarrow \\ M \times M & \overset{\mu}{\rightarrow} & M \end{array}$$

ausdrücken.

Die Beispiele 1. bis 4. sind assoziativ. Das Beispiel 5. ist es nicht:

$$\mu(1,\mu(1,2)) = 2\,, \quad \mu(\mu(1,1),2) = 3\,.$$

Sei jetzt $(M,\circ)$ eine Menge mit einer assoziativen Verknüpfung. In der Tat ist es jetzt sinnvoll, die Verknüpfung in der Form $a \circ b$ zu schreiben, da $(a \circ b) \circ c = a \circ (b \circ c)$ etc. gilt und man sich das Setzen von Klammern sparen kann.

Definition 2.4. *Ein Element $e \in M$ heißt neutrales Element der Verknüpfung, wenn es die Identitäten*

$$e \circ a = a\,, \quad a \circ e = a$$

erfüllt.

Als Diagramm geschrieben:

$$\begin{array}{ccc} M & \overset{\cdots \times e}{\rightarrow} & M \times M \\ & \mathtt{id} \searrow & \circ \downarrow \\ & & M \end{array} \quad , \qquad \begin{array}{ccc} M & \overset{e \times \cdots}{\rightarrow} & M \times M \\ & \mathtt{id} \searrow & \circ \downarrow \\ & & M \end{array} \quad .$$

Die Mengen $\mathbb{N}, \mathbb{Z}, \mathbb{Q}, \mathbb{R}$ haben ein neutrales Element $e := 1$ bezüglich der Multiplikation. Die Mengen $\mathbb{Z}, \mathbb{Q}, \mathbb{R}$ haben ein neutrales Element $e := 0$ bezüglich der Addition.

Im Beispiel 2. ist das neutrale Element $e = (0, \ldots, 0)^t \in \mathbb{R}^k$.

Die Matrix $E_n \in \mathtt{Mat}(n, n, R)$ ist im Beispiel 3.das neutrale Element bezüglich $\bullet$.

Im Beispiel 4. ist $\mathtt{id}_A \in \mathtt{Hom}(A, A)$ das neutrale Element.

Sei nun $(M, \circ, e)$ eine Menge mit Verknüpfung und neutralem Element. Sei $a \in M$.

Definition 2.5. *Ein Linksinverses $b \in M$ von a ist ein Element, welches $b \circ a = e$ erfüllt.*

Lemma 2.6. *Wir nehmen an, daß $(M, \circ, e)$ eine Menge mit Verknüpfung und neutralem Element ist, in welcher jedes Element ein Linksinverses besitzt. Dann gelten folgende Aussagen:*

1. *Ist $a, b \in M$ und $b \circ a = e$, dann gilt auch $a \circ b = e$.*
2. *Das Linksinverse von a ist eindeutig durch a bestimmt (wir schreiben a^{-1}).*
3. *Es gilt $(a^{-1})^{-1} = a$.*
4. *Es gibt kein weiteres von e verschiedenes neutrales Element in M.*

Proof. Sei $a \circ b = e$. Dann gilt für ein Linksinverses c von b daß

$$a \circ b = c \circ b \circ a \circ b = c \circ b = e \ .$$

Dies zeigt 1.

Sei nun b' ein weiteres Linksinverses von a. Dann gilt $b' \circ a = e = b \circ a$, also nach Anwenden von $\cdots \circ b$ auch $b' = b$. Dies zeigt 2.

Es gilt $(a^{-1})^{-1} \circ a^{-1} = e = a \circ a^{-1}$, also 3.

Sei e' ein weiteres neutrales Element. Dann gilt $e' = e^{-1} = e$. □

Definition 2.7. *Eine Gruppe $(M, \circ, e)$ ist eine Menge M mit assoziativer Verknüpfung $\circ$ und neutralem Element e derart, daß jedes Element von M ein Linksinverses besitzt.*

In Diagrammen kann man die Existenz des Inversen auch beschreiben, nämlich dadurch, daß eine Abbildung $I : M \to M$ existiere mit

$$\begin{array}{ccc} M & \overset{(I,\mathtt{id}_M)}{\to} & M \times M \\ \downarrow & & \circ \downarrow \\ * & \overset{e}{\to} & M \end{array} .$$

In der letzten Zeile betrachten wir die Angabe eines Elementes $e \in M$ als gleichwertig mit der Angabe einer Abbildung von der einpunktigen Menge $*$ nach M. Die weiteren Eigenschaften des Inversen sehen in Diagrammform so aus:

$$\begin{array}{ccc} M & \overset{(\mathtt{id}_M,I)}{\to} & M \times M \\ \downarrow & & \circ \downarrow \\ * & \overset{e}{\to} & M \end{array} \quad , \quad I \circ I = \mathtt{id}_M .$$

Lemma 2.8. *Sei $(M, \circ, e)$ eine Gruppe und $N \subset M$ eine nichtleere Teilmenge, welche abgeschlossen unter $\circ$ und der Bildung des Inversen ist. Dann ist $e \in N$ und $(N, \circ, e)$ eine Gruppe.*

Proof. Es bleibt nur zu zeigen, daß $e \in N$ gilt. Nun hat aber N mindestens ein Element $n \in N$. Dann gilt $n^{-1} \in N$ und auch $e = n \circ n^{-1} \in N$. □
Wir sagen, daß $N \subset M$ eine Untergruppe ist.

2.1.2 Beispiele von Gruppen - Untergruppen von Zahlbereichen

Die Menge $(\mathbb{Z}, +, 0)$ ist eine Gruppe (die additive Gruppe von $\mathbb{Z}$). In der Tat ist das inverse Element von n durch $-n$ gegeben.

Analog sind $(\mathbb{Q}, +, 0)$ und $(\mathbb{R}, +, 0)$ Gruppen (die additiven Gruppen von $\mathbb{Q}$ und $\mathbb{R}$).

Die Menge $(\mathbb{R}^k, +, (0, \ldots, 0))$ ist eine Gruppe. Das Inverse von $x \in \mathbb{R}^k$ ist $-x$.

Sei $\mathbb{Q}^* := \mathbb{Q} \setminus \{0\}$ und $\mathbb{R}^* := \mathbb{R} \setminus \{0\}$. Die Menge $(\mathbb{Q}^*, *, 1)$ ist eine Gruppe (die multiplikative Gruppe von $\mathbb{Q}$). Das Inverse von x ist x^{-1}. Analog ist $(\mathbb{R}^*, *, 1)$ die multiplikative Gruppe von $\mathbb{R}$.

2.1.3 Gruppen von Matrizen

Definition 2.9. *Die Menge der invertierbaren $n \times n$-Matrizen*

$$GL(n,\mathbb{R}) := \{A \in \mathtt{Mat}(n,n,\mathbb{R}) | A \text{ ist invertierbar}\}$$

ist die allgemeine lineare Gruppe (mit der Verknüpfung $\bullet$ und E_n als neutralem Element).

Die Menge $P(n,\mathbb{R})$ der invertierbaren oberen $n \times n$-Dreiecksmatrizen ist eine Untergruppe von $GL(n,\mathbb{R})$ (siehe Lemma 1.20).

Sei

$$N(n,\mathbb{R}) := \{A \in \mathtt{Mat}(n,n,\mathbb{R}) | j \leq i \Rightarrow A_{i,j} = 0\ .\}$$

Dies sind also obere Dreiecksmatrizen, deren Diagonalelemente verschwinden.

Definition 2.10. *Eine Matrix $X \in \mathtt{Mat}(n,n,\mathbb{R})$, für welche ein $k \in \mathbb{N}$ existiert mit $X^k = 0$, heißt nilpotent (k-ter Stufe).*

Die Elemente aus $N(n,\mathbb{R})$ sind nilpotent. In der Tat gilt für $X \in N(n,\mathbb{R})$ die Gleichung $X^n = 0$. Man rechnet induktiv nach, daß $j \leq i+l$ das Verschwinden $(X^l)_{i,j} = 0$ impliziert.

Wir schreiben 1 für E_n. Sei $X \in \mathtt{Mat}(n,n,\mathbb{R})$ nilpotent.

Lemma 2.11. *Die Matrix $1+X$ ist invertierbar.*

Proof. Das Inverse von $1+X$ ist durch die (endliche) Summe (Neumannreihe)

$$(1+X)^{-1} = \sum_{l=0}^{\infty} (-1)^l X^l$$

gegeben. In der Tat gilt

$$\begin{aligned}(1+X) \bullet \sum_{l=0}^{\infty} (-1)^l X^l &= \sum_{l=0}^{\infty} (-1)^l X^l + \sum_{l=0}^{\infty} (-1)^l X^{l+1} \\ &= \sum_{l=0}^{\infty} (-1)^l X^l - \sum_{l=1}^{\infty} (-1)^l X^l \\ &= 1\end{aligned}$$

□

Wir definieren

$$U(n,\mathbb{R}) := \{1+X | X \in N(n,\mathbb{R})\}\ .$$

Dann ist $U(n,\mathbb{R})$ eine Untergruppe von $P(n,\mathbb{R}) \subset GL(n,\mathbb{R})$.

Definition 2.12. *Eine Matrix $A \in \mathtt{Mat}(n,n,\mathbb{R})$ heißt Permutationsmatrix, wenn $A_{i,j} \in \{0,1\}$ gilt und in jeder Zeile und Spalte genau eine* 1 *vorkommt.*

Sei $W_n \subset \mathtt{Mat}(n,n,\mathbb{R})$ die Menge der Permutationsmatrizen. Offensichtlich gilt $1 \in W_n$. Die Matrizen $P_n(i,j)$ (siehe 1.3.2) gehören zu W_n.

Lemma 2.13. *W_n ist eine Untergruppe von $GL(n,\mathbb{R})$.*

Proof. Kommt später 2.16. □

2.1.4 Permutationen - die Gruppen S_n

Sei A eine Menge.

Definition 2.14. *Die Menge der Bijektionen von A*

$$S(A) := \{ f \in \mathtt{Hom}(A,A) | A \textit{ ist bijektiv} \}$$

ist die symmetrische Gruppe von A. Insbesondere ist die symmetrische Gruppe von $\overline{n} = \{1,\ldots,n\}$ die n-te Permutationsgruppe

$$S_n := S(\overline{n}) \ .$$

Sei $A = \{a,b,c,d\}$. Dann kann ein Element $f \in S(A)$ durch eine Wertetabelle dargestellt werden.

x	a	b	c	d
$f(x)$	b	c	a	d

Eine kürzere Schreibweise für dieses Element ist die Zyklendarstellung

$$(abc) \ .$$

Damit wird beschrieben, daß für diese Permutation $a \mapsto b$, $b \mapsto c$ und $c \mapsto a$ gilt, sowie daß d fest bleibt. Hier ein weiteres Beispiel $h \in S_9$:

x	1	2	3	4	5	6	7	8	9
$h(x)$	2	4	6	5	1	7	3	8	9

In Zyklen ist dieses Element das Produkt

$$(1,2,4,5) \circ (3,6,7) \ .$$

Wir betrachten $(1,2),(2,3) \in S_3$. Dann gilt

$$(1,2) \circ (2,3) = (1,2,3), \quad (2,3) \circ (1,2) = (1,3,2) \ .$$

Das Ergebnis der Komposition kann also durchaus von der Reihenfolge abhängen.

Ein Zyklus der Länge 2 wird auch Transposition genannt.

Sei $A = B \cup C$ und $B \cap C = \emptyset$. Wir haben eine natürliche Einbettung $i_B : S(B) \hookrightarrow S(A)$, welche $\psi \in S(B)$ durch $\mathtt{id}_C$ auf A fortsetzt.

Lemma 2.15. *Sei A eine endliche Menge. In $S(A)$ kann jedes Element als Produkt von Transpositionen dargestellt werden.*

Proof. Wir sehen zuerst ein, daß jedes Element aus $S(A)$ durch ein Produkt von Zyklen dargestellt werden kann. Wir führen eine Induktion nach der Anzahl der Elemente von A durch. Ist $|A| = 1$, dann ist die Aussage offensichtlich. Wir führen nun den Schritt $n \Rightarrow n+1$ vor. Möge also $|A| = n+1$ gelten. Sei $\psi \in S(A)$. Sei $a \in A$. Wir definieren eine Folge $a_i = \psi^i(a)$, $i = 1,2,\ldots$. Sei $l > 0$ minimal mit $a_l = a$. Dann betrachten wir den Zyklus $\phi := (a_0, a_1, \ldots, a_{l-1})$. Sei $C := \{a_0, \ldots, a_{l-1}\}$ und $B := A \setminus C$. Dann gilt offensichtlich $\psi(C) = C$ und $\psi(B) = B$. Es gilt $|B| \leq n$. Nach Induktionsvoraussetzung läßt sich $\psi_{|B} \in S(B)$ als Produkt von Zyklen $z_1 \circ \cdots \circ z_r$ darstellen. Dann gilt aber die Zerlegung in Zyklen von ψ

$$\psi = i_B(z_1) \circ \cdots \circ i_B(z_r) \circ \phi \ .$$

Wir zeigen nun, daß sich ein Zyklus als Produkt von Transpositionen schreiben läßt. Wir führen wieder Induktion nach der Länge durch. Die wichtige Rechnung ist dabei

$$(a_1, \ldots, a_n) \circ (a_{n+1}, a_n) = (a_1, \ldots, a_{n+1}) \ .$$

□

Wir wollen nun S_n mit W_n vergleichen und Lemma 2.13 beweisen.

Lemma 2.16. *Es gibt zueinander inverse Bijektionen $\sigma : W_n \to S_n$ und $A : S_n \to W_n$, welche mit den Multiplikationen verträglich sind. W_n ist insbesondere eine Gruppe.*

Proof. Wir betrachten die Elemente $e_i \in \mathbb{R}^n$ (siehe 1.18). Ist $A \in W_n$, dann gibt es genau ein $\sigma(A) \in S_n$ derart, daß $Ae_{\sigma(A)(i)} = e_i$ ist. Folglich gilt $A_{i,j} = \delta_{i,\sigma(A)(j)}$.

Wir erhalten eine Abbildung $\sigma : W_n \to S_n$. Umgekehrt, wenn $\phi \in S_n$ ist, dann definieren wir $A(\phi) \in W_n$ durch

$$A(\phi)_{i,j} = \delta_{i,\phi(j)} \ .$$

Wir erhalten eine Abbildung $A : S_n \to W_n$.

Es gilt $A \circ \sigma = \mathtt{id}$ und $\sigma \circ A = \mathtt{id}$. In der Tat gilt $A(\sigma(B))_{i,j} = \delta_{i,\sigma(B)(j)} = B_{i,j}$ und $\delta_{i,\sigma(A(\phi))(j)} = A(\phi)_{i,j} = \delta_{i,\phi(j)}$, also $\sigma(A(\phi)) = \phi$.

Die Abbildungen sind in der Tat mit den Verknüpfungen verträglich. Es gilt

$$A(\phi \circ \psi) = A(\phi) \circ A(\psi) \ .$$

In der Tat gilt

$$\delta_{i,(\phi\circ\psi)(j)} = A(\phi \circ \psi)_{i,j} \ .$$

Weiter gilt

$$\begin{aligned}
(A(\phi) \circ A(\psi))_{i,j} &= \sum_{k=1}^{n} A(\phi)_{i,k} A(\psi)_{k,j} \\
&= \sum_{k=1}^{n} \delta_{i,\phi(k)} \delta_{k,\psi(j)} \\
&= \sum_{k=1}^{n} \delta_{i,k} \delta_{\phi^{-1}(k),\psi(j)} \\
&= \delta_{\phi^{-1}(i),\psi(j)} \\
&= \delta_{i,(\phi\circ\psi)(j)} \ .
\end{aligned}$$

Das Bild $A(S_n) = W_n$ ist also abgeschlossen unter der Multiplikation. Da auch $A(1) = 1 \in W_n$ ist, und $A(\phi) \circ A(\phi^{-1}) = A(1) = 1$ gilt, ist W_n eine Gruppe. □

Im Lemma haben wir einen Isomorphismus der Gruppen W_n und S_n konstruiert. Wir werden diese Begrifflichkeit später (siehe 2.1.6) genauer studieren.

2.1.5 Die Gruppen $\mathbb{Z}/n\mathbb{Z}$ und $\mathbb{Z}/n\mathbb{Z}^*$

Wir beschreiben die allgemeine Konstruktion der Bildung der Menge der Äquivalenzklassen. Sei $(X,\sim)$ eine Menge mit einer Äquivalenzrelation. Zur Erinnerung, $\sim$ ist reflexiv, symmetrisch und transitiv.

Definition 2.17. *Sei $x \in X$. Dann heißt die Menge*

$$[x]_\sim := \{y \in X | x \sim y\}$$

die Äquivalenzklasse von x.

Es gilt offensichtlich $x \in [x]$ und damit $X = \cup_{x\in X}[x]$.

Lemma 2.18. *Seien $x,y \in X$. Dann gilt entweder $[x] = [y]$ oder $[x] \cap [y] = \emptyset$.*

Proof. Sei $[x] \cap [y] \neq 0$. Es genügt zu zeigen, daß $[x] \subset [y]$ gilt. Sei $b \in [x] \cap [y]$. Sei $a \in [x]$. Dann gilt $a \sim x$ und $b \sim x$, also $a \sim b$. Wegen $b \sim y$ gilt auch $a \sim y$, also $a \in [y]$. □

Wenn $y \in [x]$ ist, so gilt $[y] = [x]$.

Definition 2.19. *Wir definieren die Menge der Äquivalenzklassen*

$$X/\sim := \{[x] | x \in X\} \subset \mathcal{P}(X)\ .$$

Wir sehen also, daß die Menge X in eine disjunkte Vereinigung von Äquivalenzklassen zerfällt.

Wir benutzen diese Konstruktion, um weitere Beispiele von Gruppen zu konstruieren. Wir fixieren $n \in \mathbb{Z}$. Wir definieren auf $\mathbb{Z}$ eine Relation $\sim_n$ durch

$$x \sim_n \Leftrightarrow n|(x-y)$$

($n|x-y$ ist die Aussage, daß n die Differenz $x-y$ teilt).

Lemma 2.20. *$\sim_n$ ist eine Äquivalenzrelation.*

Proof. Nachrechnen ! □

Definition 2.21. *Wir definieren die Menge*

$$\mathbb{Z}/n\mathbb{Z} := \mathbb{Z}/\sim_n \ .$$

Offensichtlich gilt $\mathbb{Z}/n\mathbb{Z} := \{[0],[1],[2],\ldots,[n-1]\}$.

Wir definieren nun eine Addition $+ : \mathbb{Z}/n\mathbb{Z} \times \mathbb{Z}/n\mathbb{Z} \to \mathbb{Z}/n\mathbb{Z}$. Dazu beobachten wir folgendes. Seien $I,J \in \mathbb{Z}/n\mathbb{Z}$ und $a \in I, b \in J$. Dann hängt $[a+b]$ nur von I und J ab. In der Tat, wenn $a' \in I$ und $b' \in J$ gilt, so gibt es $k,l \in \mathbb{Z}$ mit $a-a' = kn$ und $b-b' = ln$. Es folgt $(a+b)-(a'+b') = (k+l)n$.

Definition 2.22. *Wir definieren* $+ : \mathbb{Z}/n\mathbb{Z} \times \mathbb{Z}/n\mathbb{Z} \to \mathbb{Z}/n\mathbb{Z}$ *durch* $I+J := [a+b]$, *wobei* $a \in I$ *und* $b \in J$ *beliebige Vertreter sind.*

Hier ist die Additionstabelle von $\mathbb{Z}/6\mathbb{Z}$.

	[0]	[1]	[2]	[3]	[4]	[5]
[0]	[0]	[1]	[2]	[3]	[4]	[5]
[1]	[1]	[2]	[3]	[4]	[5]	[0]
[2]	[2]	[3]	[4]	[5]	[0]	[1]
[3]	[3]	[4]	[5]	[0]	[1]	[2]
[4]	[4]	[5]	[0]	[1]	[2]	[3]
[5]	[5]	[0]	[1]	[2]	[3]	[4]

.

Lemma 2.23. $(\mathbb{Z}/n\mathbb{Z}, +, [0])$ *ist eine Gruppe.*

Proof. Wir zeigen zuerst, daß $+$ assoziativ ist. Nachrechnen ! Danach sehen wir ein, daß $[0]$ das neutrale Element ist. Nachrechnen ! Schließlich sehen wir, daß $[-i]$ das Inverse von $[i]$ ist. □

Wir definieren nun eine weitere Verknüpfung $*$ auf $\mathbb{Z}/n\mathbb{Z}$, wobei wir das gleiche Prinzip wie bei der Addition anwenden. Wir beobachten, daß für $a \in I$ und $b \in J$ die Klasse von $[ab]$ nur von I und J abhängt. In der Tat, wenn $a' + kn = a$ und $b' + ln = b$, dann gilt $ab = a'b' + (a'l + b'k)n$.

Definition 2.24. *Wir definieren die Multiplikation* $* : \mathbb{Z}/n\mathbb{Z} \times \mathbb{Z}/n\mathbb{Z} \to \mathbb{Z}/n\mathbb{Z}$ *durch* $I * J = [ab]$ *für beliebige Vertreter* $a \in I$ *und* $b \in J$.

Hier ist die Multiplikationstabelle für $\mathbb{Z}/6\mathbb{Z}$:

	[0]	[1]	[2]	[3]	[4]	[5]
[0]	[0]	[0]	[0]	[0]	[0]	[0]
[1]	[0]	[1]	[2]	[3]	[4]	[5]
[2]	[0]	[2]	[4]	[0]	[2]	[4]
[3]	[0]	[3]	[0]	[3]	[0]	[3]
[4]	[0]	[4]	[2]	[0]	[4]	[2]
[5]	[0]	[5]	[4]	[3]	[2]	[1]

.

Das neutrale Element von $\mathbb{Z}/n\mathbb{Z} \setminus \{[0]\}$ bezüglich $*$ ist $[1]$.

Lemma 2.25. $(\mathbb{Z}/n\mathbb{Z} \setminus \{[0]\}, *, [1])$ *eine Menge mit assoziativer Verknüpfung und neutralem Element.*

Proof. Nachrechnen ! □

Diese ist jedoch im allgemeinen keine Gruppe. Im Beispiel $\mathbb{Z}/6\mathbb{Z} \setminus \{[0]\}$ hat etwa $[2]$ kein inverses Element. Wir werden später (2.3.2) sehen, daß $(\mathbb{Z}/n\mathbb{Z} \setminus \{[0]\}, *, [1])$ genau dann eine Gruppe ist, wenn n eine Primzahl ist. An dieser Stelle können wir uns als Gruppe einfach die Teilmenge der Elemente herausgreifen, die ein Inverses besitzen.

Definition 2.26. *Sei* $(\mathbb{Z}/n\mathbb{Z})^* \subset \mathbb{Z}/n\mathbb{Z}$ *die Teilmenge der Elemente, die ein Linksinverses besitzen. Die Gruppe* $((\mathbb{Z}/n\mathbb{Z})^*, *, [1])$ *ist die Gruppe der Einheiten im Ring* $\mathbb{Z}/n\mathbb{Z}$*.*

Im Beispiel ist

$$(\mathbb{Z}/6\mathbb{Z})^* = \{[1], [5]\}$$

und es gilt $[5] * [5] = 1$, also $[5]^{-1} = [5]$.

Definition 2.27. *Wir definieren die Eulersche* φ*-Funktion* $\varphi : \mathbb{N} \to \mathbb{N}$ *durch*

$$\varphi(n) := |(\mathbb{Z}/n\mathbb{Z})^*| \, .$$

Wir haben $\varphi(1) = 1$, $\varphi(6) = 2$ und nach obiger Bemerkung für eine Primzahl p, daß $\varphi(p) = p - 1$. Die Struktur der φ-Funktion wird in der Zahlentheorie untersucht.

2.1.6 Symmetrien und Gruppen

Mathematische Objekte kommen meist mit einem angepaßten Begriff von Abbildungen oder Transformationen.

Betrachtet man einfach Mengen A, B, dann sind Transformationen gerade Abbildungen $A \to B$. Wenn man auf den Mengen zusätzliche Strukturen betrachtet, dann sollen die entsprechenden Transformationen damit verträglich sein. Seien etwa die Elemente von A und B mit der Farbe weiß oder rot versehen, daß heißt wir haben Abbildungen $a : A \to \{w, r\}$ und $b : B \to \{w, r\}$ gegeben. Eine strukturerhaltende Abbildung $f : A \to B$ ist dann eine, welche weisse auf weisse und rote auf rote Elemente abbildet. In Formeln kann diese Bedingung als

$$f^* b := b \circ f = a$$

beschrieben werden. Als Diagramm kann diese Gleichung in der Form

$$\begin{array}{ccc} A & \xrightarrow{f} & B \\ & {\scriptstyle a}\searrow \quad \swarrow{\scriptstyle b} & \\ & \{w, r\} & \end{array}$$

geschrieben werden.

Eine andere Struktur auf einer Menge ist etwa die Festlegung eines Abstandes zwischen ihren Punkten. Ein Abstand auf A ist eine Funktion $d_A : A \times A \to [0, \infty)$ (mit zusätzlichen hier nicht interessanten Eigenschaften). Eine abstandserhaltende Abbildung (Isometrie) $f : A \to B$ sollte dann

$$d_B(f(a), f(a')) = d_A(a, a')$$

erfüllen. Also, der Abstand der Bilder ist gleich dem Abstand der Urbilder.

Auch eine Verknüpfung ist eine zu betrachtende Struktur. Eine verknüpfungserhaltende Abbildung $f : A \to B$ sollte dann

$$f(a) \circ_B f(a') = f(a \circ_A a')$$

erfüllen. Also, das Bild von verknüpften Elementen ist gleich der Verknüpfung der Bilder der Elemente. Als Diagramm geschrieben sieht diese Gleichung so aus:

$$\begin{array}{ccc} A \times A & \overset{f \times f}{\to} & B \times B \\ {\scriptstyle \circ_A}\downarrow & & {\scriptstyle \circ_B}\downarrow \\ A & \xrightarrow{f} & B \end{array} .$$

Wir werden viele weitere solche Beispiele sehen. Insbesondere kann man die strukturerhaltenden Abbildungen $A \to A$ betrachten. Dies ist eine Teilmenge

$$\mathtt{Hom}_{Struktur}(A,A) \subset \mathtt{Hom}_{Menge}(A,A) \ .$$

Struktur steht hier für Farbe, Abstand, Verknüpfung etc. Die Menge der invertierbaren strukturerhaltenden Abbildungen

$$\mathtt{Aut}_{Struktur}(A) := \{ f \in \mathtt{Hom}_{Struktur}(A,A) | \, f \text{ ist invertierbar} \}$$

ist eine Gruppe, nämlich die Gruppe der Symmetrien von A und ihrer Struktur.

Hier ist ein Beispiel. Wir betrachten die Gruppe $(\mathbb{Z}/n\mathbb{Z}, +, [0])$. Sei $[j] \in Z/n\mathbb{Z}$ und $f_{[j]} : \mathbb{Z}/n\mathbb{Z} \to \mathbb{Z}/n\mathbb{Z}$ durch $f_{[j]}([k]) := [j][k]$ gegeben.

Lemma 2.28. *Die Abbildung $f_{[j]}$ erhält die Gruppenstruktur.*

Proof.

$$\begin{aligned} f_{[j]}([l]+[k]) &= [j]([k]+[l]) \\ &= [j][k+l] \\ &= [j(k+l)] \\ &= [jk+jl] \\ &= [jk]+[jl] \\ &= [j][k]+[j][l] \\ &= f_{[j]}([k]) + f_{[j]}([l]) \ . \end{aligned}$$

□

Wenn $[j] \in (\mathbb{Z}/n\mathbb{Z})^*$, dann ist $f_{[j]}$ eine Bijektion, also $f_{[j]} \in \mathtt{Aut}_{groups}(\mathbb{Z}/n\mathbb{Z})$.

2.1.7 Abelsche Gruppen

Eine typische Eigenschaft der additiven und multiplikativen Gruppen von Zahlbereichen ist, daß das Ergebnis der Verknüpfung unabhängig von der Reihenfolge ist. Solche Gruppe nennen wir abelsch.

Definition 2.29. *Ein Gruppe $(G, \circ, 1)$ heißt abelsch, falls in ihr die Identität $a \circ b = b \circ a$ gilt.*

In Diagrammen kann man das so schreiben:

$$\begin{array}{ccc} G\times G & \stackrel{T}{\to} & G\times G \\ \circ\downarrow & & \circ\downarrow \\ G & = & G \end{array} ,$$

wobei $T : G\times G \to G\times G$ die Transposition $T(a,b) := (b,a)$ ist.

Weitere Beispiele für abelsche Gruppen sind $\mathbb{Z}/n\mathbb{Z}$, $(\mathbb{Z}/n\mathbb{Z})^*$ und $\mathbb{R}^n$.

Die Gruppen S_n für $n \geq 3$ und $GL(n,\mathbb{R})$ für $n \geq 2$ sind nicht abelsch.

2.2 Ringe

2.2.1 Die Ringaxiome

In vielen Beispielen sind auf ein und derselben Menge zwei Verknüpfungen erklärt. So ist etwa auf $\mathbb{N},\mathbb{Z},\mathbb{Q},\mathbb{R},\mathbb{Z}/n\mathbb{Z}$ und auf $\mathtt{Mat}(n,n,\mathbb{R})$ sowohl eine Addition $+$ als auch eine Multiplikation $*$ erklärt. Wenn die Menge mit der Addition eine abelsche Gruppe bildet und die Multiplikation mit der Addition verträglich ist (Distributivgesetz), dann spricht man von einer Ringstruktur.

Definition 2.30. *Ein Ring ist eine Menge R mit zwei assoziativen Verknüpfungen $+,* : R\times R \to R$ und neutralem Element 0 der Addition mit den folgenden Eigenschaften:*

1. *$(R,+,0)$ ist eine abelsche Gruppe.*
2. *Es gilt das Distributivgesetz:*

$$a*(b+c) = a*b+a*c\ , \quad (a+b)*c = a*c+b*c\ .$$

Das Distributivgesetz besagt, daß die Abbildungen $a*\ldots, \cdots * c : R \to R$ (Linksmultiplikation mit a bzw. Rechtsmultiplikation mit c) die Gruppenstruktur $(R,+,0)$ erhalten.

Definition 2.31. *Ein Ring mit 1 ist ein Ring mit einem neutralen Element der Multiplikation.*

Definition 2.32. *Ein kommutativer Ring ist ein Ring, in welchem die Multiplikation kommutativ ist, d.h in welchem die Relation $a*b = b*a$ gilt.*

2.2.2 Beispiele für Ringe

Es ist klar, daß $\mathbb{Z}, \mathbb{Q}, \mathbb{R}$ kommutative Ringe mit 1 sind. $\mathbb{N}$ ist kein Ring, da die natürlichen Zahlen unter der Addition keine Gruppe bilden.

Die Matrizen $\mathtt{Mat}(n,n,\mathbb{R})$ etwa bilden einen Ring mit 1, sind aber für $n \geq 2$ nicht kommutativ.

Lemma 2.33. $(\mathbb{Z}/n\mathbb{Z}, +, *, [0], [1])$ *ist ein kommutativer Ring mit* 1.

Proof. Nach der Diskussion in 2.1.5 bleibt das Distributivgesetz zu zeigen. Wir haben in Lemma 2.28 gezeigt, daß die Multiplikation von links die Gruppenstruktur erhält. In diesem Fall ist die Multiplikation kommutativ. □

2.2.3 Polynome und formale Potenzreihen

Wir betrachten den Ring der Polynome. Ein Polynom $\leq n$-ten Grades über $\mathbb{R}$ ist eine Funktion f mit der Darstellung

$$f(x) = f_n x^n + \cdots + f_0 \ .$$

Wir werden $f(x) = \sum_{i=0}^{\infty} f_i x^i$ schreiben, wobei verabredungsgemäß $f_i = 0$ für $i > n$ gesetzt wird, und die Summe hier nur über die endlich vielen Summanden mit $f_i \neq 0$ genommen wird. Ist $g(x) = g_m x^m + \cdots + g_0$ ein weiteres Polynom, so sind $f+g$ und fg Polynome vom Grade $\leq \mathtt{max}(n,m)$ und $\leq nm$. Es gilt

$$\begin{aligned} (f+g)(x) &= \sum_{i=0}^{\infty} (f_i + g_i) x^i \\ (fg)(x) &= \sum_{i=0}^{\infty} \Big(\sum_{k+l=i} f_k g_l \Big) x^i \end{aligned} \tag{3}$$

Das wesentliche an einem Polynom ist die Darstellung $f(x) = f_n x^n + \cdots + f_0$. Ein Polynom kann also als eine Abbildung $f : \mathbb{N}_0 \to \mathbb{R}$, $i \mapsto f_i$, verstanden werden, wobei höchstens endlich viele Werte nicht verschwinden. Die Identifikation mit Funktionen ist hier nur zur Motivation der Rechenoperationen benutzt worden.

Ist R irgend ein kommutativer Ring mit 1.

Definition 2.34. *Die Menge der Polynome $R[x]$ ist die Menge der Folgen $f : \mathbb{N}_0 \to \mathbb{R}$, $i \mapsto f_i$, bei denen höchstens endlich viele Werte nicht verschwinden. Die Addition und die Multiplikation wird wie in (3) erklärt. Sei $0 \in R[x]$ die Nullfolge, und $1 \in R[x]$ die Folge* $1,0,0\ldots$.

Dann gilt

Lemma 2.35. $(R[x],+,*,0,1)$ *ist ein kommutativer Ring mit* 1.

Proof. Nachrechnen! □

Jedem Polynom $f \in R[x]$ kann eine Funktion $R \to R$, $\lambda \mapsto f(\lambda)$ zugeordnet werden. Diese Zuordnung ist aber im allgemeinen nicht injektiv. Sei z.B. $R = \mathbb{Z}/2\mathbb{Z}$. Dann ist die Funktion zu $f(x) = x^2 + x$ die Nullfunktion, es gilt aber $f \neq 0$. In diesem Fall sind Polynome nicht als Funktionen zu betrachten.

Sei $R[[x]] := R^{\mathbb{N}_0}$ die Menge aller Funktionen $\mathbb{N}_0 \to R$. Die Ringoperationen auf $R[x]$ kann man auf $R[[x]]$ mit denselben Formeln ausdehnen.

Definition 2.36. $R[[x]]$ *ist der Ring der formalen Potenzreihen.*

Wir schreiben Elemente aus $R(x)$ sinnvoller Weise in der Form

$$a_0 + xa_1 + x^2a_2 \ldots \ .$$

Der Ring $R[[x]]$ ist ganz interessant. Es ist klar, daß nichtkonstante Polynome in $R[x]$ in der Regel keine Inversen haben. Anders dagegen in $R[[x]]$. So ist zum Beispiel die Potenzreihe $1-x$ invertierbar. Das Inverse ist

$$(1-x)^{-1} = 1 + x + x^2 + x^3 + \ldots \ .$$

Auf der anderen Seite definiert ein Polynom $f \in R[x]$ eine Funktionen $R \to R$ durch $\lambda \mapsto f(\lambda)$. In formale Potenzreihen kann man keine Ringelemente einsetzen und so einen Wert definieren.

2.3 Körper

2.3.1 Körperaxiome

Unter den kommutativen Ringen mit 1 sind die Körper dadurch ausgezeichnet, daß von Null verschiedene Elemente ein Inverses bezüglich der Multiplikation besitzen.

Definition 2.37. *Ein Körper* $(K,+,*,0,1)$ *ist ein kommutativer Ring mit* 1*, für welchen* $(K\setminus\{0\},*,1)$ *eine Gruppe ist.*

Beispiele für Körper sind $\mathbb{Q}$ und $\mathbb{R}$. Die ganzen Zahlen sind kein Körper.

2.3.2 Die Körper $\mathbb{Z}/p\mathbb{Z}$

Wir betrachten den Ring $\mathbb{Z}$. Die Einheiten in $\mathbb{Z}$ sind die multiplikativ invertierbaren Elemente, also 1 und -1. Die Primzahlen sind Zahlen $p\in\mathbb{N}$, welche nur von $\pm p$ und ± 1 geteilt werden. Jede andere ganze Zahl kann eindeutig als Produkt $up_1^{n_1}\dots p_r^{n_r}$ mit $u\in\{1,-1\}$, Primzahlen $0<p_1<p_2<\dots$ und $n_i\in\mathbb{N}$ dargestellt werden. Dies ist eine Aussage, die eigentlich eines Beweises bedarf. Dieser wird etwa in der elementaren Zahlentheorie erbracht. Wir werden hier diese Tatsache einfach akzeptieren.

Sei $n\in\mathbb{N}$.

Lemma 2.38. *Der Ring* $\mathbb{Z}/n\mathbb{Z}$ *ist genau dann ein Körper, wenn n eine Primzahl ist.*

Proof. Ist n keine Primzahl, so gibt es eine nichttriviale Zerlegung $n=pq$. Dann sind $[p],[q]\in\mathbb{Z}/n\mathbb{Z}\setminus\{[0]\}$ und $[p]*[q]=[pq]=[n]=[0]$. Da auch $[0]*[q]=0$ gilt, kann $[q]$ kein multiplikatives Inverses haben. (Zur Veranschaulichung betrachten wir die Multiplokationstabelle von $\mathbb{Z}/6\mathbb{Z}$ und sehen etwa, daß $[2][3]=0$.)

Sei nun $[n]$ eine Primzahl. Sei $r\in\{1,\dots,n-1\}$. Dann ist $[r]\neq 0$ Wir zeigen, daß $[r]*\cdots:\mathbb{Z}/n\mathbb{Z}\to\mathbb{Z}/n\mathbb{Z}$ injektiv ist. Wäre diese Abbildung nicht injektiv, also etwa $[r][i]=[r][j]$ mit $[i]\neq[j]$, dann auch $[r][s]=0$ mit $[s]=[i]-[j]\neq 0$. Dann gilt $rs=ln$ für ein geeignetes $l\in\mathbb{Z}$. Die Primzahl n muß als Primfaktor von r oder s auftreten, woraus $[r]=0$ oder $[s]=0$ folgt. Wir hatten aber $[r]$ und $[s]$ von Null verschieden gewählt.

Wir haben nun gesehen, daß $[r] * \cdots : \mathbb{Z}/n\mathbb{Z} \to \mathbb{Z}/n\mathbb{Z}$ injektiv ist. Da $\mathbb{Z}/n\mathbb{Z}$ endlich ist, folgt auch die Surjektivität. Folglich kommt $[1]$ im Bild von $[r] * \ldots$ vor. Damit besitzt $[r]$ ein Rechtsinverses. Dieses ist wegen der Kommutativität der Multiplikation auch Linksinvers. □

Als Beispiel betrachten wir die Multiplikationstabelle des Körpers $\mathbb{Z}/5\mathbb{Z}$.

	0	1	2	3	4
0	0	0	0	0	0
1	0	1	2	3	4
2	0	2	4	1	3
3	0	3	1	4	2
4	0	4	3	2	1

.

Definition 2.39. *Sei p eine Primzahl. Dann schreiben wir*

$$\mathbb{F}_p := \mathbb{Z}/p\mathbb{Z} \, .$$

Es gilt $|\mathbb{F}_p| = p$. Wegen $\mathbb{F}_p^* = \mathbb{F}_p \setminus \{0\}$ gilt $\varphi(p) = p - 1$ (siehe 2.27).

2.3.3 Quadratische Zahlörper

Als ein weiteres Beispiel betrachten wir die Menge der Zahlen

$$\mathbb{Q}[\sqrt{3}] = \{a + b\sqrt{3} | a, b \in \mathbb{Q}\} \subset \mathbb{R}$$

mit den von $\mathbb{R}$ induzierten Operationen.

Lemma 2.40. $\mathbb{Q}[\sqrt{3}]$ *ist Körper.*

Proof. Wir überprüfen zuerst durch einfaches Nachrechnen, daß diese Menge abgeschlossen unter den Operationen $+$ und $*$ ist. Als nächstes zeigen wir, daß $\mathbb{Q}[\sqrt{3}]$ ein Ring ist. In der Tat folgt aus $x \in \mathbb{Q}[\sqrt{3}]$ auch $-x \in \mathbb{Q}[\sqrt{3}]$. Schließlich schreiben wir

$$\frac{1}{a + b\sqrt{3}} = \frac{a}{a^2 - 3b^2} - \sqrt{3}\frac{b}{a^2 - 3b^2} \, ,$$

wobei wir ausnutzen, daß $\sqrt{3} \in \mathbb{R} \setminus \mathbb{Q}$ ist. Damit ist nämlich der Nenner nur dann Null, wenn $a = b = 0$ gilt.

Wäre $\sqrt{3} \in \mathbb{Q}$, dann würde $\sqrt{3} = \frac{p}{q}$ für ganze Zahlen p, q gelten. Damit wäre $3q^2 = p^2$. Auf der rechten (linken) Seite kommt der Primfaktor 3 in einer ungeraden (geraden) Vielfachheit vor. Dies widerspricht der eindeutigen Primfaktorzerlegung in $\mathbb{Z}$. □

Nach diesem Prinzip kann man viele Körper zwischen $\mathbb{Q}$ und $\mathbb{R}$ konstruieren.

Insbesondere kann man für jedes $a \in \mathbb{Q}$ mit $a > 0$ einen kleinsten Körper $K \subset \mathbb{R}$ konstruieren, in welchem die Gleichung $x^2 = a$ eine Lösung hat.

2.3.4 Die komplexen Zahlen

In diesem Abschnitt wollen wir einen Körper $\mathbb{C}$ konstruieren, in welchem die Gleichung $x^2 = -1$ eine Lösung hat. Wir betrachten dazu die Matrix

$$I := \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \in \mathtt{Mat}(2,2,\mathbb{R}) \ .$$

Diese Matrix erfüllt $I^2 = -E_2$. Wir betrachten nun die Teilmenge

$$\mathbb{C} := \{aE_2 + bI | a, b \in \mathbb{R}\} \subset \mathtt{Mat}(2,2,\mathbb{R}) \ .$$

mit den vom Ring $(\mathtt{Mat}(2,2,\mathbb{R}), +, \bullet, 0_2, E_2)$ induzierten Operationen.

Definition 2.41. *$(\mathbb{C}, +, \bullet, 0_2, E_2)$ ist ein Körper.*

Proof. Wir rechnen nach, daß $\mathbb{C}$ unter $+$ und $\bullet$ abgeschlossen ist. Weiter überprüfen wir, daß $\mathbb{C}$ ein Ring ist. Schließlich sehen wir ein, daß $\mathbb{C}$ ein kommutativer Ring mit 1 ist. Wir rechnen nach, daß das Inverse von $aE_2 + bI$ durch

$$(aE_2 + bI)^{-1} = \frac{aE_2 - bI}{a^2 + b^2}$$

gegeben ist. □

Wir haben

$$z = aE_2 + bI = \begin{pmatrix} a & -b \\ b & a \end{pmatrix} \ .$$

Folglich werden die Zahlen $a, b \in \mathbb{R}$ durch das Element z eindeutig bestimmt.

Definition 2.42. *Ist $z = aE_2 + bI$ eine komplexe Zahl, so heißen die Zahlen $\Re(z) := a$ und $\Im(z) := b$ der Real- und der Imaginärteil von z.*

Wir können z mit dem Paar $(a,b)^t \in \mathbb{R}^2$ identifizieren. Die Addition in $\mathbb{R}^2$ geht dabei in die Addition von $\mathbb{C}$ über. Die Multiplikation ist durch (Nachrechnen)

$$(a,b)t(a',b')^t = (aa' - bb', ab' + a'b)^t$$

gegeben. Das Inverse berechnet sich durch $((a,b)^t)^{-1} = (\frac{a}{a^2+b^2}, \frac{-b}{a^2+b^2})^t$. Wir können den Körper $\mathbb{R}$ als Unterkörper von $\mathbb{C}$ verstehen, indem wir $a \in \mathbb{R}$ mit aE_2 (bzw. $(a,0)^t \in \mathbb{R}^2$) identifiziert. Wir werden die komplexe Zahl $aE_2 + bI$ oft auch als $a + bi$ schreiben.

In Anlehnung an die Notation $\mathbb{Q}[\sqrt{3}]$ kann man auch $\mathbb{C} = \mathbb{R}[\sqrt{-1}]$ schreiben und behaupten, daß $i = \sqrt{-1}$ gilt.

Im folgenden betrachten wir der Körper $\mathbb{C}$ als eine Menge mit Struktur. Wir können dann die Automorphismen $\mathtt{Aut}_{fields}(\mathbb{C})$ betrachten.

Lemma 2.43. *Die Abbildung $a + ib \mapsto \overline{a+ib} := a - ib$ ist ein Automorphismus von $\mathbb{C}$.*

Proof. Nachrechnen ! □

Wir nennen diesen Automorphismus die komplexe Konjugation.

Definition 2.44. *Wir definieren die Abbildung (den Betrag)*

$$|\ldots| : \mathbb{C} \to \mathbb{R}$$

durch

$$|z|^2 := z\overline{z}\,.$$

Es gelten die folgenden weiteren Eigenschaften.

Lemma 2.45.

1. *Wenn $z \in \mathbb{R}$, so gilt $\overline{z} = z$.*
2. *Es gilt $\overline{\overline{z}} = z$.*
3. *Es gilt $|zz'| = |z||z'|$.*
4. *Es gilt $|z|^{-1} = |z^{-1}|$.*
5. *Für $a,b \in \mathbb{R}$ gilt $|a+bi|^2 = a^2 + b^2$.*

Proof. Nachrechnen! □

Wir wollen hier das Rechnen mit komplexen Zahlen, insbesondere die Multiplikation, geometrisch veranschaulichen. Dazu greifen wir auf einige Elemente der euklidischen Geometrie zurück, welche aus der Schule bekannt sein sollten, in diesem Kurs jedoch erst später behandelt werden.

Auf dem Raum $\mathbb{R}^2$ betrachten wir das Skalarprodukt

$$< \ldots, \cdots >: \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R} ,$$

welches durch die Formel $< (u_1,u_2)^t, (v_1,v_2)^t >:= u_1v_1 + u_2v_2$ gegeben wird. Mit Hilfe des Skalarproduktes definieren wir die Norm

$$\| \ldots \| : \mathbb{R}^2 \to \mathbb{R} , \|x\| := \sqrt{< x,x >} .$$

Um Winkel zu messen, betrachten wir die Äquivalenzrelation auf $\mathbb{R}$, welche durch $x \sim y \Leftrightarrow x - y \in 2\pi\mathbb{Z}$ gegeben ist. Wir schreiben $\mathbb{R}/2\pi\mathbb{Z} := \mathbb{R}/\sim$. Elemente aus $\mathbb{R}/2\pi\mathbb{Z}$ können vertreterweise addiert und mit ganzen Zahlen multipliziert werden. Den Winkel $\phi \in \mathbb{R}/2\pi\mathbb{Z}$ zwischen zwei nichttrivialen Elementen $u, v \in \mathbb{R}^2$ definieren wir durch

$$\cos(\phi) = \frac{< u,v >}{\|u\|\|v\|} .$$

Zur Erinnerung, eine komplexe Zahl $a+bi$ ist eine reelle Matrix. Wir untersuchen nun die Abbildung $(a+bi)\bullet : \mathbb{R}^2 \to \mathbb{R}^2$. Es gilt $(a+0i) = \mathtt{diag}(a,a)$. Folglich wirkt $(a+0i)\bullet$ durch skalare Multiplikation mit a.

Wir untersuchen nun die Veränderung der Länge. Sei $v = (v_1,v_2)^t \in \mathbb{R}^2$.

$$\begin{aligned} \|(a+bi)(v_1,v_2)^t\|^2 &= \|(av_1 - bv_2, bv_1 + av_2)^t\|^2 \\ &= a^2v_1^2 - 2abv_1v_2 + b^2v_2^2 + b^2v_1^2 + 2abv_1v_2 + a^2v_2^2 \\ &= (a^2+b^2)(v_1^2+v_2^2) \\ &= |a+bi|^2\|v\|^2 \end{aligned}$$

Folglich gilt für $z \in \mathbb{C}$ und $v \in \mathbb{R}^2$, daß

$$\|z \bullet v\| = |z|\|v\| .$$

Wenn $|z| = 1$, dann erhält $z\bullet$ die Länge und ist eine Drehung. Dazu berechnen wir den Winkel ϕ zwischen v und zv. Sei $\|v\| = 1$. Es gilt

$$\begin{aligned}\cos(\phi) &= < zv, v > \\ &= < (av_1 - bv_2, bv_1 + av_2)^t, (v_1, v_2)^t > \\ &= av_1^2 - bv_1v_2 + bv_1v_2 + av_2^2 \\ &= a\end{aligned}$$

Wenn $z = a + bi$ mit $|z| = 1$ ist, dann dreht also $z\bullet$ jeden Vektor um den Winkel $\cos(\phi) = a$.

Sei nun $0 \neq z = (a + bi) \in \mathbb{C}$. Dann schreiben wir $z = |z|\frac{z}{|z|}$. Folglich wirkt z als Hintereinanderausführung einer Drehung um den Winkel ϕ (welcher durch $\cos(\phi) = \frac{a}{\sqrt{a^2+b^2}}$ gegeben wird) und einer Streckung um $\sqrt{a^2 + b^2}$.

Sind $x, y \in \mathbb{C}$, dann ist die Operation von xy auf $\mathbb{R}^2$ durch die Hintereinanderausführung der Operationen von y und x gegeben. Seien ϕ und ψ die Drehwinkel von x, y. Dann ist xy die Komposition einer Drehung um den Winkel $\phi + \psi$ und einer Streckung um den Faktor $|x||y|$.

Sei $n \in \mathbb{N}$ gegeben. Die Lösungen der Gleichung $x^n = 1$ heißen n-te Einheitswurzeln. Ist $\xi \in \mathbb{C}$ eine n-te Einheitswurzel, dann gilt $|\xi| = 1$. Ist ϕ der Drehwinkel von ξ, so muß $n\phi = 0$ (in $\mathbb{R}/2\pi\mathbb{Z}$) gelten. Wir erhalten folgende n verschiedene Möglichkeiten

$$\phi_m = [2\pi\frac{m}{n}] \ , m = 0, \ldots, n-1 \ .$$

Die n-ten Einheitswurzeln in $\mathbb{C}$ sind also durch

$$\xi_m := \cos(2\pi\frac{m}{n}) + i\sin(2\pi\frac{m}{n}), \ , m = 0, \ldots, n-1$$

gegeben. Die 4ten Einheitswurzeln sind beispielsweise

$$1, i, -1, -i \ .$$

Die dritten Einheitswurzeln sind

$$(1, -\frac{1}{2} + i\frac{\sqrt{3}}{2}, -\frac{1}{2} - i\frac{\sqrt{3}}{2}) \ .$$

Mit Hilfe der Polynomdivision kann man zeigen, daß

$$x^n - 1 = \prod_{m=0}^{n-1}(x - \xi_m)$$

gilt. Wir sagen, daß das Polynom $x^n - 1$ in Linearfaktoren zerfällt.

Die komplexen Zahlen haben die Eigenschaft, daß jedes Polynom

$$p(x) = x^n + a_{n-1}x^{n-1} + \cdots + a_0$$

mit $a_i \in \mathbb{C}$ in Linearfaktoren zerfällt, d.h. es gibt Zahlen $\lambda_1, \ldots, \lambda_n \in \mathbb{C}$ mit

$$p(x) = \prod_{i=1}^{n} (x - \lambda_i) \; .$$

Dies ist im wesentlichen die Aussage des Fundamentalsatzes der Algebra, den wir allerdings an dieser Stellen noch nicht beweisen wollen.

Wir erinnern an die Formel für die Nullstellen eines quadratischen Polynoms $f(x) = x^2 + px + q$:

$$x_\pm := -\frac{p}{2} \pm \sqrt{\frac{p^2}{4} - q} \; . \tag{4}$$

Arbeitet man mit reellen Zahlen, dann ist diese Formel nur im Fall $\frac{p^2}{4} - q \geq 0$ sinnvoll. Betrachtet man komplexe Zahlen, dann ist diese Formel auch ohne diese Voraussetzung gültig. Wenn $p, q \in \mathbb{R}$ und $\frac{p^2}{4} - q < 0$, so sind

$$x_\pm = -\frac{p}{2} \pm i\sqrt{-\frac{p^2}{4} + q}$$

die komplexen Nullstellen von f. In der Tat liefert die Formel (4) die Nullstellen für beliebige quadratische Polynome mit komplexen Koeffizienten. Dazu müssen wir allerdings die Wurzel aus komplexen Zahlen ziehen können.

Im geometrischen Bild ist das ganz einfach. Sei ϕ der Drehwinkel von z. Der Betrag einer Wurzel u aus z ist $|u| = \sqrt{|z|}$, während der Drehwinkel ψ von u der Gleichung $2\psi = \phi$ genügt. Diese hat zwei Lösungen. Sei $\tilde{\phi} \in [0, 2\pi)$ ein Repräsentant von ϕ. Dann sind die Lösungen durch $\frac{1}{2}\tilde{\phi}$ und $\frac{1}{2}(\tilde{\phi} + 2\pi)$ repräsentiert.

Rechnerisch erhält man die Wurzeln aus z wie folgt. Die Zahl $z + |z|$ dreht um den halben Winkel von z. Man muß also nur noch normieren und bekommt

$$\pm\sqrt{z} = \pm \frac{\sqrt{|z|}}{|z + |z||} (z + |z|) \; .$$

2.3.5 Rationale Funktionen

Sei jetzt K ein Körper und $K[x]$ der Ring der Polynome über K. Es ist klar, daß $K[x]$ kein Körper ist. Zum Beispiel besitzt $f(x) = x$ kein Inverses. Es gibt jedoch eine Methode, diesen Ring zu einem Körper zu erweitern. Diese Methode kennen wir sicher vom Übergang von $\mathbb{Z}$ nach $\mathbb{Q}$. Daß die folgende Prozedur funktioniert, liegt daran, daß $fg = 0$ zur Folge hat, daß $f = 0$ oder $g = 0$ gilt (wir sagen, der Ring $K[x]$ hat keine Nullteiler). Dann betrachten wir die Menge der Paare $(f,g) \in K[x] \times (K[x] \setminus \{0\})$. Auf dieser Menge definieren wir die Relation $(f,g) \sim (f',g')$ falls es $h,h' \in K[x] \setminus \{0\}$ gibt mit $(fh,gh) = (f'h',g'h')$. Sei $QK[x]$ die Menge der Äquivalenzklassen. Die Klasse von (f,g) schreiben wir in der Form $\frac{f}{g}$. Wir erkären ferner die Operationen

$$\begin{aligned} \frac{f}{g} + \frac{f'}{g'} &= \frac{fg' + f'g}{gg'} \\ \frac{f}{g}\frac{f'}{g'} &= \frac{ff'}{gg'} . \end{aligned}$$

Wir betrachten $K[x] \subset QK[x]$,indem wir f mit $(f,1)$ identifizieren. Beachte, daß $(f,1) \sim (f',1)$ die Gleichheit $f = f'$ nach sich zieht.

Lemma 2.46. *Diese Operationen sind wohldefiniert (Vertreterunabhängigkeit). Weiter ist* $(QK[x],+,*,1,0)$ *ein Körper (der Quotientenkörper von* $K[x]$*).*

Proof. Nachrechnen. □

Der Körper $QK[x]$ ist der Körper der rationalen Funktionen.

2.3.6 Quaternionen

Zur Erinnerung, ein Körper ist ein kommutativer Ring mit Eins, in welchem jedes von Null verschiedene Element ein multiplikatives Inverses hat. Die Bedingung "kommutativ" kann man weglassen und kommt so zum Begriff des Schiefkörpers.

Definition 2.47. *Ein Schiefkörper ist ein Ring mit Eins* $(S,+,*,0,1)$*, für welchen* $(S \setminus \{0\},*,1)$ *eine Gruppe ist.*

Wir wollen hier als Beispiel die Quaternionen kennenlernen.

Dazu betrachten wird die komplexen 2×2-Matrizen

$$I := \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}, J := \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, K := \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix} .$$

Wir lassen das Multiplikationszeichen weg und schreiben $1 := E_2$. Man rechnet leicht nach, daß

$$I^2 = J^2 = K^2 = -1, \quad IJ = -JI = K, \quad KI = -IK = J, \quad JK = -KJ = I$$

gilt. Wir definieren die Teilmenge

$$\mathbb{H} := \{a1 + bI + cJ + dK | a, b, c, d \in \mathbb{R}\} \subset \mathtt{Mat}(2, 2, \mathbb{C}) .$$

Es gilt

$$a1 + bI + cJ + dK = \begin{pmatrix} a + bi & c + id \\ -c + id & a - bi \end{pmatrix} .$$

Das heißt, die Darstellung eines Quaternions in der Form $a1 + bI + cJ + dK$ ist eindeutig. Als Menge können wir $\mathbb{H}$ mit $\mathbb{R}^4$ identifizieren.

Lemma 2.48. *Mit der Matrizenaddition und Multiplikation ist $\mathbb{H}$ ein Ring mit Eins.*

Wir können $\mathbb{R}$ und $\mathbb{C}$ als Unterring von $\mathbb{H}$ verstehen, wenn man $a + bi$ mit $a1 + bI$ identifiziert. Dies werden wir im folgenden immer machen:

$$\mathbb{R} \subset \mathbb{C} \subset \mathbb{H} .$$

Wir definieren die Konjugation $\overline{\ldots} : \mathbb{H} \to \mathbb{H}$ durch

$$\overline{a1 + bI + cJ + dK} := a1 - bI - cJ - dK$$

und $|q|^2 := q\bar{q}$. Wir rechnen nach, daß für $q = a1 + bI + cJ + dK$ gilt:

$$q\bar{q} = a^2 + b^2 + c^2 + d^2 .$$

Lemma 2.49. *Die Quaternionen sind ein Schiefkörper.*

Proof. Das Inverse von $q \in \mathbb{H} \setminus \{0\}$ ist durch $\frac{\bar{q}}{|q|^2}$ gegeben. □

Definition 2.50. *Der Realteil und der Imaginärteil von $q = a + bI + cJ + dK$ ist durch $\Re(q) := a$ und $\Im(q) = bI + cJ + dK$ definiert.*

Wir können die imaginären Quaternionen mit $\mathbb{R}^3$ via $(b,c,d)^t \mapsto bI + cJ + dK$ identifizieren. Wir können nun ein Produkt auf $\mathbb{R}^3$ durch

$$\times : \mathbb{R}^3 \times \mathbb{R}^3 \to \mathbb{H} \times \mathbb{H} \stackrel{prod}{\to} \mathbb{H} \stackrel{\Im}{\to} \mathbb{R}^3$$

definieren. Wir berechnen

$$\begin{aligned}(b,c,d)^t \times (e,f,g)^t &= \Im((bI+cJ+dK)(eI+fJ+gK)) \\ &= \Im(-be-cf-dk+(cg-df)I+(de-bg)J+(bf-ce)K) \\ &= (cg-df, de-bg, bf-ce)^t\end{aligned}$$

Das ist also gerade das aus der Schule bekannte Kreuzprodukt. Diese Produkt ist nicht assoziativ. In der Tat gilt etwa

$$e_1 \times (e_1 \times e_2) = -e_2 \ , \quad (e_1 \times e_1) \times e_2 = 0 \ .$$