1 Matrizen und Determinanten

1.1 Lineare Gleichungssysteme

Ein Beispiel eines Gleichungssystems ist

$$\begin{aligned} x^4 - 3x^2 + y^2 + xy + 6 &= 0 \\ 4x^2 - 3y^2 - 3xy - 10 &= 0 \end{aligned}$$

Eine Lösung dieses Systems besteht aus zwei reelle Zahlen x, y die eingesetzt die beiden Gleichungen erfüllen. Eine Lösung ist also ein *geordnetes Paar* (x, y) von Zahlen.

Um eine Lösung von $(*)$ zu finden subtrahieren wir die zweite Gleichung von der ersten und vertauschen die Seiten:

$$\begin{aligned} 0 &= x^4 - 7x^2 + 4y^2 + 4xy + 16 \\ &= (x^2 - 4)^2 + x^2 + 4y^2 + 4xy \\ &= (x^2 - 4)^2 + (x + 2y)^2 \end{aligned}$$

Also muss $x^2 - 4 = 0 = x + 2y$ gelten und wir erhalten

$$(x, y) = (2, -1), \quad \text{oder} \quad (x, y) = (-2, 1).$$

Aber keines dieser Paare löst auch nur eine der beiden Anfangsgleichungen. Der Grund dafür ist, dass wir im ersten Schritt unseres "Lösungsverfahrens" Information verschenkt, d.h. Bedingungen ignoriert und damit die Lösungsmenge vergrößert haben. In Wirklichkeit hat das System keine reelle Lösungen.

Im Allgemeinen ist es sehr schwierig zu entscheiden, ob ein Gleichungssystem lösbar ist und ggf. die Menge aller Lösungen zu bestimmen.

Wir untersuchen im Folgenden nur *lineare Gleichungssysteme*. Eine lineare Gleichung ist eine Gleichung der Form

$$a_1x_1 + \ldots + a_nx_n = b.$$

Dabei sind $a_1 \ldots a_n$, b gegebene $n+1$ Zahlen und $x_1 \ldots x_n$ gesuchte n Zahlen.

Das Adjektiv "linear" bezieht sich darauf, dass nicht etwa Quadrate, Produkte oder noch kompliziertere Funktionen der Unbekannten $x_1 \ldots x_n$ in der Gleichung vorkommen. Hat man nur wenige Unbekannte so bezeichnet man diese auch mit $x, y, z, \ldots$ statt $x_1, x_2, x_3, \ldots$

Lineare Gleichungssysteme sind mehrere solcher Gleichungen, z.B.

$$\begin{aligned} 3x - 2y + 5z &= 2 \\ y + 9z &= 0 \end{aligned}$$

Ein lineares Gleichungssystem ist also ein Gleichungssystem der Form

$$\begin{array}{ccccccccc} a_{11}x_1 & + & a_{12}x_2 & + & \ldots & + & a_{1n}x_n & = & b_1 \\ a_{21}x_1 & + & a_{22}x_2 & + & \ldots & + & a_{2n}x_n & = & b_2 \\ \vdots & & & & & & \vdots & & \vdots \\ a_{m1}x_1 & + & a_{m2}x_2 & + & \ldots & + & a_{mn}x_n & = & b_m \end{array}$$

Dabei sind die a_{ij} und die b_i als bekannt vorausgesetzt. Man hat also m Gleichungen für die n Unbekannten $x_1, \ldots, x_n$. Eine Lösung des Systems ist ein geordnetes n-Tupel $(x_1, \ldots, x_n)$ von Zahlen, das alle m Gleichungen erfüllt.

Im Gegensatz zu beliebigen Gleichungssystemen gibt es für lineare Gleichungen gibt ein allgemeines Verfahren, einen Algorithmus, der

- zur Lösung führt, wenn es eine gibt,
- alle Lösungen liefert, wenn es mehrere gibt,
- gegebenenfalls meldet, daß es keine Lösung gibt.

Dieses Verfahren heißt Gaußscher Algorithmus. Wir führen ihn zunächst an zwei Beispielen vor.

1.1.1 Beispiel:

$$\begin{array}{rcrcr} 5x & + & 7y & = & 3 \\ 2x & + & 3y & = & -1 \end{array}$$

Wir multiplizieren die erste Zeile mit $-\frac{2}{5}$ und addieren sie zur zweiten und erhalten das lineare Gleichungssystem

$$\begin{array}{rcrcr} 5x & + & 7y & = & 3 \\ & & \frac{1}{5}y & = & \frac{-11}{5} \end{array}$$

Die Lösung der zweiten Gleichung ist $y = -11$. Setzen wir diesen Wert in die erste Gleichung ein und lösen diese nach x auf, so erhalten wir $x = 16$. Das Gleichungssystem hat also genau eine Lösung, nämlich $(x, y) = (16, -11)$.

1.1.2 Beispiel:

$$(1) \qquad \begin{array}{rcrcrcr} 2x_1 & + & x_2 & + & 3x_3 & = & 0 \\ 4x_1 & - & x_2 & + & x_3 & = & 10 \\ 3x_1 & + & 2x_2 & + & x_3 & = & 7 \end{array} \qquad \begin{array}{l} \\ -2 \times \text{ erste Zeile} \\ -\frac{3}{2} \times \text{ erste Zeile} \end{array}$$

$$(2) \qquad \begin{array}{rcrcrcr} 2x_1 & + & x_2 & + & 3x_3 & = & 0 \\ & - & 3x_2 & - & 5x_3 & = & 10 \\ & & \frac{1}{2}x_2 & - & \frac{7}{2}x_3 & = & 7 \end{array} \qquad \begin{array}{l} \\ \\ +\frac{1}{6} \times \text{ zweite Zeile} \end{array}$$

$$(3) \qquad \begin{array}{rcrcrcr} 2x_1 & + & x_2 & + & 3x_3 & = & 0 \\ & - & 3x_2 & - & 5x_3 & = & 10 \\ & & & & -\frac{13}{3}x_3 & = & \frac{26}{3} \end{array}$$

Das dritte Gleichungssystem ist sehr leicht zu lösen. Aus der letzten Gleichung folgt $x_3 = -2$. Einsetzen in die vorletzte Gleichung liefert $x_2 = 0$, und Einsetzen der gefundenen Werte in die erste Gleichung des dritten Systems liefert $x_1 = 3$. Also hat das System die einzige Lösung $(x_1, x_2, x_3) = (3, 0, -2)$.

Beachten Sie, dass man bei den Umformungen keine Information verliert: Addiert man z.B. im Gleichungssystem (2) zur zweiten Zeile das Doppelte der erste Zeile und zur dritten Zeile das $\frac{3}{2}$-fache der ersten Zeile so erhält man das ursprüngliche Gleichungssystem (1) zurück. Alle Umformungen sind umkehrbar.

Wir beschreiben nun den Gaußschen Algorithmus im allgemeinen. Gegeben sei ein lineares Gleichungssystem

$$\begin{array}{ccccccccc} a_{11}x_1 & + & a_{12}x_2 & + & \dots & + & a_{1n}x_n & = & b_1 \\ a_{21}x_1 & + & a_{22}x_2 & + & \dots & + & a_{2n}x_n & = & b_2 \\ \vdots & & & & & & \vdots & & \vdots \\ a_{m1}x_1 & + & a_{m2}x_2 & + & \dots & + & a_{mn}x_n & = & b_m \end{array}$$

Wir nehmen an, dass x_1 wirklich vorkommt, d.h. eines der a_{i1} ist $\neq 0$. Sonst nummerieren wir die Unbekannten um. Nach Vertauschung von zwei Zeilen können wir annehmen, dass sogar $a_{11} \neq 0$. Dann addieren wir von den folgenden Zeilen ein Vielfaches der ersten, so dass jeweils x_1 verschwindet. Genauer: wir addieren zur i-ten Zeile das $-\frac{a_{i1}}{a_{11}}$-fache der ersten. Wir erhalten ein neues, zum ersten System äquivalentes, in dem x_1 nur noch in der ersten Zeile vorkommt:

$$\begin{array}{ccccccccc} a_{11}x_1 & + & a_{12}x_2 & + & \dots & + & a_{1n}x_n & = & b_1 \\ & & a'_{22}x_2 & + & \dots & + & a'_{2n}x_n & = & b'_2 \\ & & \vdots & & & & \vdots & & \vdots \\ & & a'_{m2}x_2 & + & \dots & + & a'_{mn}x_n & = & b'_m \end{array}$$

Noch einmal die vorgenommen Umformungen:

- Vertauschung von zwei Zeilen,
- Addition eines Vielfachen einer Zeile zu einer anderen.

Bei den weiteren Umformungen des Systems betrachten wir nur noch die Zeile ab der zweiten, die erste Zeile führen wir unverändert mit. Wir betrachten in den folgenden Zeilen die Unbekannte mit dem kleinsten noch vorkommenden Index. In der Regel wird das x_2 sein, es könnte aber auch sein, dass alle Terme in denen x_2 vorkam, bei den Umformungen verschwunden sind. Nehmen wir an, x_j wäre diese Unbekannte. Nach Zeilenvertauschung können wir wieder annehmen, dass x_j in der zweiten Zeile wirklich vorkommt. Dann addieren wir wie eben zu den Zeilen $3, \dots, m$ jeweils ein geeignetes Vielfaches der zweiten Zeile, so dass x_j in keiner der Zeilen $3, \dots, m$ mehr vorkommt.

Dieses Verfahren setzen wir fort, bis in den nachfolgenden Zeilen gar keine Unbekannten mehr stehen. Das ist insbesondere dann der Fall, wenn es keine nachfolgenden Zeilen mehr gibt. Es kann aber auch sein, dass noch Gleichungen da sind, die aber auf der linken Seite keine Unbekannten mehr haben also von der Form

$$0 \quad = \quad b$$

für irgendeine rechte Seite b sind. Das System ist dann in *Zeilenstufenform.*

1.1.3 Definition: Ein lineares Gleichungssystem

$$
(*) \qquad \begin{array}{ccccccccc}
a_{11}x_1 & + & a_{12}x_2 & + & \ldots & + & a_{1n}x_n & = & b_1 \\
a_{21}x_1 & + & a_{22}x_2 & + & \ldots & + & a_{2n}x_n & = & b_2 \\
\vdots & & & & & & \vdots & & \vdots \\
a_{m1}x_1 & + & a_{m2}x_2 & + & \ldots & + & a_{mn}x_n & = & b_m
\end{array}
$$

ist in Zeilenstufenform, falls für jeden Zeilenindex $i = 1, 2, \ldots, m$ ein Index $j \in \{1, \ldots, n\}$ existiert, so daß

(1) $a_{ij} \neq 0$ und $a_{ik} = 0 \quad \forall\ k = 1, \ldots, j-1$

(2) $a_{kl} = 0 \quad \forall\ k > i, l \leq j$.

Ein System in Zeilenstufenform sieht also so aus:

Da der Index j in (1) von i abhängt schreiben wir auch j_i dafür. Wegen (2) gilt:

$$j_1 < j_2 < \ldots < j_r \leq n.$$

Wir wollen jetzt untersuchen wann das System $(*)$ in Zeilenstufenform lösbar ist und beschreiben wie die Lösungsmenge gegebenenfalls aussieht.

1. Fall: $r < m$ und mindestens eine der Zahlen $b_{r+1}, b_{r+2}, \ldots, b_m$ ist $\neq 0$. Dann besitzt das System $(*)$ keine Lösung.

2. Fall: $b_{r+1} = \ldots = b_m = 0$ oder $r = m$.

Dann können wir die letzten $m - r$ Zeilen ignorieren, sie enthalten keine Information.

Das Gleichungssystem ist in diesem Fall lösbar.

Fall 2, (a): $r = n$. Dann sind alle Stufen von der Breite 1, also $j_i = i$ für $i = 1, 2, \ldots, n$. Die letzte Gleichung ist $a_{nn}x_n = b_n$ und wir erhalten daraus einen eindeutig bestimmten Wert für x_n. Diesen setzen wir in die vorletzte Gleichung

$$a_{n-1\,n-1}x_{n-1} + a_{n-1\,n}x_n = b_{n-1}$$

ein und erhalten daraus den Wert für x_{n-1}. So fortfahrend sehen wir, daß das Gleichungssystem eine eindeutig bestimmte Lösung $(x_1, x_2, \ldots, x_n)$ besitzt.

Fall 2, (b): $r < n$. Dann gibt es wenigstens eine breitere Stufe d.h. es gibt mindestens eine Variable x_j die nicht am Anfang einer Zeile steht. Eine solche Variable nennnen wir "freie Variable". Davon gibt es offenbar genau $n - r$ Stück. Wir können sie beliebig wählen und für jede solche Wahl erhält man eine eindeutig bestimmte Lösung.

Wir führen den Algorithmus noch einmal an einem Beispiel vor. Gegeben sei das Gleichungssystem

$$(*) \qquad \begin{array}{ccccccc} & & x_2 & + & 2x_3 & = & 9 \\ 3x_1 & + & 4x_2 & + & 5x_3 & = & 9 \\ 6x_1 & + & 7x_2 & + & 8x_3 & = & 9 \\ 9x_1 & + & 9x_2 & + & 9x_3 & = & 9. \end{array}$$

Wir benutzten dafür die abkürzende Schreibweise

$$\begin{array}{ccc|c} 0 & 1 & 2 & 9 \\ 3 & 4 & 5 & 9 \\ 6 & 7 & 8 & 9 \\ 9 & 9 & 9 & 9 \end{array}$$

Als ersten Schritt vertauschen wir die 1. und 2. Zeile:

$$\begin{array}{ccc|cl} 3 & 4 & 5 & 9 & \\ 0 & 1 & 2 & 9 & \\ 6 & 7 & 8 & 9 & \qquad \text{-2}\times \text{ 1. Zeile} \\ 9 & 9 & 9 & 9 & \qquad \text{-3}\times \text{ 1. Zeile} \end{array}$$

$$\Updownarrow$$

$$\begin{array}{cccc|l} 3 & 4 & 5 & 9 & \\ 0 & 1 & 2 & 9 & \\ 0 & -1 & -2 & -9 & \qquad 1\times \text{ 2. Zeile} \\ 0 & -3 & -6 & -18 & \qquad 3\times \text{ 2. Zeile} \end{array}$$

$$\Updownarrow$$

$$\begin{array}{ccc|c} 3 & 4 & 5 & 9 \\ 0 & 1 & 2 & 9 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 9 \end{array}$$

Also ist das System nicht lösbar.

Wenn wir stattdessen die letzte Zeile von $(*)$ durch die Gleichung

$$(*) \qquad 9x_1 + 9x_2 + 9x_3 = 0$$

ersetzen so stossen wir am Ende auf das Diagramm

$$\begin{array}{ccc|c} 3 & 4 & 5 & 9 \\ 0 & 1 & 2 & 9 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}$$

gestoßen, d.h. in diesem Fall ist das System äquivalent zu dem Gleichungssystem

$$(**) \qquad \begin{array}{ccccccc} 3x_1 & + & 4x_2 & + & 5x_3 & = & 9 \\ & & x_2 & + & 2x_3 & = & 9 \end{array}$$

Wir erhalten unendlich viele Lösungen, da wir x_3 frei wählen können, d.h. zu jeder Wahl von x_3 gibt es genau eine Lösung. Als Lösungsmenge von $(**)$ erhalten wir

$$\{(t-9, 9-2t, t) \mid t \in \mathbb{R}\}.$$

1.1.4 Definition: 1. Ein lineares Gleichungssystem der Form

$$\begin{array}{ccccccc} a_{11}x_1 & + & \dots & + & a_{1n}x_n & = & 0 \\ \vdots & & & & & & \vdots \\ a_{m1}x_1 & + & \dots & + & a_{mn}x_n & = & 0 \end{array}$$

heißt homogen.

2. Setzt man in einem beliebigen linearen Gleichungssystem

$$\begin{array}{ccccccc} a_{11}x_1 & + & \dots & + & a_{1n}x_n & = & b_1 \\ \vdots & & & & & & \vdots \\ a_{m1}x_1 & + & \dots & + & a_{mn}x_n & = & b_n \end{array}$$

die rechte Seite $= 0$, so erhält man ein neues lineares Gleichungssystem, dass das zugehörige homogene System genannt wird.

1.1.5 Bemerkung: Hat das homogene lineare Gleichungssystem

$$\begin{array}{ccccccc} a_{11}x_1 & + & \ldots & + & a_{1n}x_n & = & 0 \\ \vdots & & & & \vdots & & \\ a_{m1}x_1 & + & \ldots & + & a_{mn}x_n & = & 0 \end{array}$$

mehr Unbekannte als Gleichungen, d.h. ist $m < n$, so besitzt das System eine nicht-triviale Lösung

$$(x_1, \ldots, x_n) \neq (0, \ldots, 0).$$

Beweis: Wir bringen das System mit dem Gaußalgorithmus auf Zeilenstufenform mit r nicht-trivialen Zeilen. Da $r \leq m < n$ gibt es mindestens eine "freie Variablen". Sei x_j eine solche. Nach den obigen Überlegungen gibt es eine Lösung $(x_1, \ldots, x_n)$ des Systems bei der wir x_j beliebig vorgeben können (also etwa $x_j = 1$). $\square$

1.2 Matrizenrechnung

1.2.1 Im Folgenden sei K entweder die Menge der rationalen Zahlen $\mathbb{Q}$, oder die Menge der reelle Zahlen $\mathbb{R}$.

1.2.2 Definition: Eine $m \times n$-Matrix (oder Matrix vom Typ $m \times n$) über K besteht aus mn Zahlen in K die in einem rechteckigen Schema der Form

$$A = \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{pmatrix}.$$

angeordnet sind. Die Elemente $a_{ij} \in K$ $1 \leq i \leq m, 1 \leq j \leq n$ heißen Komponenten oder Einträge der Matrix A. Wir schreiben abkürzend

$$A = (a_{ij})_{\substack{i=1,\ldots,m \\ j=1,\ldots,n}}$$

oder auch nur $A = (a_{ij})$, wenn der Typ feststeht. Matrizen vom Typ $n \times n$ heißen quadratische Matrizen.

Die $1 \times n$–Matrizen (bzw. $m \times 1$–Matrizen) heißen Zeilen- (bzw. Spalten-) vektoren. Sie haben die Form

$$z = (a_1\ a_2\ \ldots\ a_n) \quad (\text{bzw. } s = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_m \end{pmatrix})$$

Mit $M(m \times n, K)$ bezeichnen wir die Menge aller $m \times n$-Matrizen mit Komponenten aus K. Wir definieren die Summe $A + B$ zweier Matrizen $A = (a_{ij}), B = (b_{ij} \in M(m \times n, K)$ und die *skalare Multiplikation* λA einer Matrix $A = (a_{ij})$ mit einer Zahl λ wie folgt:

$$\begin{pmatrix} a_{11} & \dots & a_{1n} \\ a_{21} & & a_{2n} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix} + \begin{pmatrix} b_{11} & \dots & b_{1n} \\ b_{21} & \dots & b_{2n} \\ \vdots & & \vdots \\ b_{m1} & \dots & b_{mn} \end{pmatrix} := \begin{pmatrix} a_{11}+b_{11} & \dots & a_{1n}+b_{1n} \\ a_{21}+b_{21} & \dots & a_{2n}+b_{2n} \\ \vdots & & \vdots \\ a_{m1}+b_{m1} & \dots & a_{mn}+b_{mn} \end{pmatrix}$$

$$\lambda \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix} := \begin{pmatrix} \lambda a_{11} & \dots & \lambda a_{1n} \\ \vdots & & \vdots \\ \lambda a_{m1} & \dots & \lambda a_{mn} \end{pmatrix}$$

1.2.3 Beispiel:

$$\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix} + \begin{pmatrix} 1 & 0 & 3 \\ 0 & 5 & 2 \end{pmatrix} = \begin{pmatrix} 2 & 2 & 6 \\ 4 & 10 & 8 \end{pmatrix}, \qquad 2\begin{pmatrix} 2 & 0 \\ 3 & 1 \end{pmatrix} = \begin{pmatrix} 4 & 0 \\ 6 & 2 \end{pmatrix}.$$

Für $A = (a_{ij}) \in M(m \times n, K)$ setzen wir $-A\colon = (-a_{ij})$. Für $A, B \in M(m \times n, K)$ schreiben wir auch $A - B$ anstelle von $A + (-B)$. Offenbar ist $A - A = 0$, wobei 0 hier die $m \times n$-Matrix bezeichnet, deren Einträge alle $= 0$ sind. Diese heißt Nullmatrix. Um Mehrdeutigkeiten zu vermeiden müssten wir eigentlich $0_{m\times n}$ schreiben. Das ist aber meist zu umständlich. Aus dem Kontext wird klar werden welche Nullmatrix wir jeweils mit 0 meinen.

1.2.4 Satz: (Rechenregeln für die Addition und skalare Multiplikation von Matrizen).

(i) $\lambda(A + B) = \lambda A + \lambda B$ für alle $\lambda \in K$ und $A \in M(m \times n, K)$, $B \in M(m \times n, K)$.

(ii) $(\lambda + \mu)A = \lambda A + \mu A$ für alle $\lambda, \mu \in K$ und $A \in M(m \times n, K)$.

(iii) $(\lambda\mu)A = \lambda(\mu A)$ für alle $\lambda, \mu \in K$ und $A \in M(m \times n, K)$.

Man kann zwei Matrizen vom Typ $m\times n$ und. $n\times r$ miteinander multiplizieren und erhält dabei eine $m\times r$ Matrix als Produkt: Ist $A = (a_{ij}) \in M(m\times n, K)$

und $B = (b_{jk}) \in M(n \times r, K)$, so definieren wir das Produkt $AB \in M(m \times r, K)$ durch

$$A \cdot B = \left(\sum_{j=1}^{n} a_{ij} b_{jk} \right)_{\substack{i=1,\dots,m \\ k=1,\dots,r}}$$

1.2.5 Beispiele: (a) $(2\ 3\ 0\ 5\ 1) \begin{pmatrix} 0 \\ 1 \\ 4 \\ 1 \\ 0 \end{pmatrix} = 8$. Allgemeiner ist das Produkt eines Zeilenvektors $x = (x_1, \dots, x_n) \in \mathbb{R}^n = M(1 \times n, \mathbb{R})$ mit einem Spaltenvektor $y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \in M(n \times 1, \mathbb{R})$ eine 1×1 Matrix, d.h. eine Zahl, nämlich $xy = \sum_{i=1}^{n} x_i y_i$.

(b) Sei $A = \begin{pmatrix} 4 & 3 & 0 & 1 & 2 \\ 2 & 1 & 4 & 0 & 1 \\ 0 & 0 & 4 & 1 & 0 \\ 2 & 0 & 1 & 0 & 4 \end{pmatrix}$ und $B = \begin{pmatrix} 1 & 2 & 0 \\ 0 & 4 & 2 \\ 3 & 0 & 1 \\ 1 & 1 & 3 \\ 0 & 0 & 4 \end{pmatrix}$. Dann ist $A \cdot B = \begin{pmatrix} 5 & 21 & 17 \\ 14 & 8 & 10 \\ 13 & 1 & 7 \\ 5 & 4 & 17 \end{pmatrix}$. Der Ausdruck $B \cdot A$ ist dagegen nicht definiert.

Sind $a_1, \dots, a_m \in M(1 \times n, K)$ die m Zeilen von $A = (a_{ij}) \in M(m \times n, K)$ aufgefasst als Zeilenvektoren und $b_1, \dots, b_r \in M(n \times 1, K)$ die r Spaltenvektoren von $B = (b_{jk}) \in M(n \times r, K)$ so gilt: $A \cdot B = (a_i \cdot b_j)_{\substack{i=1,\dots,m \\ j=1,\dots,r}}$.

Die quadratische Matrix

$$E_n = (\delta_{ij}) = \begin{pmatrix} 1 & 0 & \dots & & 0 \\ 0 & 1 & & & \\ \vdots & & \ddots & & \vdots \\ & & & & 0 \\ 0 & & \dots & 0 & 1 \end{pmatrix} \in M(n \times n, K)$$

heisst $n \times n$-Einheitsmatrix. Für die Einträge von E_n gilt also $\delta_{ij} = 1$ falls gilt $i = j$ und $\delta_{ij} = 0$ sonst.

1.2.6 Satz: (Rechenregeln für die Matrizenmultiplikation).

(i) $(A_1+A_2)B = A_1B+A_2B, \quad A(B_1+B_2) = AB_1+AB_2 \qquad \forall\, A, A_1, A_2, \in M(m \times n, K), B, B_1, B_2 \in M(n \times r, K)$;

(ii) $\alpha(AB) = (\alpha A)B = A(\alpha B) \qquad \forall\, \alpha \in K,\, A \in M(m \times n, K),\, B \in M(n \times r, K)$;

(iii) $A(BC) = (AB)C \qquad \forall\, A \in M(m \times n, K), B \in M(n \times r, K), C \in M(r \times s, K)$;

(iv) $E_m A = AE_n = A \qquad \forall\, A \in M(m \times n, K)$.

1.2.7 Bemerkung: Die Matrizenmultiplikation ist i.a. nicht kommutativ. Für $A \in M(m\times n, K)$, $B \in M(n\times r, K)$ kann man zwar AB bilden nicht aber BA falls $r \neq m$. Wenn $r = n$ ist und $m \neq n$, so ist AB eine $m \times m$-Matrix während BA eine $n\times n$-Matrix ist. Selbst wenn A und B beide $n\times n$-Matrizen sind gilt i.a. $AB \neq BA$. Beispielsweise ist

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

1.2.8 Bemerkung: Eine nützliche Beobachtung, die im Folgenden noch häufiger benutzt wird, ist, dass die j-te Spalte der Matrix

$$A = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix} \in M(m \times n)$$

gleich dem Produkt

$$\begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix} \begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix} \leftarrow j\text{-te Stelle}$$

ist, d.h. die j-te Spalte von A ist $= Ae_j$ wobei

$$e_j : = \begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix} \leftarrow j\text{-te Stelle.}$$

Entsprechend sind $e_1^* A, \ldots, e_m^* A$ die Zeilen von A, wobei $(e_1^*, \ldots, e_m^*)$ die Zeilen von E_m sind.

Im Folgenden betrachten wir quadratische Matrizen.

1.2.9 Definition: Eine $n \times n$-Matrix A heißt invertierbar, wenn es eine $n \times n$-Matrix B gibt, so daß

$$AB = BA = E_n.$$

Das nächste Lemma zeigt, dass B in diesem Fall eindeutig bestimmt ist. Es wird mit A^{-1} bezeichnet (d.h. $A^{-1} : = B$) und heißt inverse Matrix oder die Inverse von A.

1.2.10 Lemma: Wenn es zu $A \in M(n \times n, K)$ zwei Matrizen $B, C \in M(n \times n, K)$ gibt mit $BA = AC = E_n$, dann ist A invertierbar und es ist $B = C = A^{-1}$.

Beweis: $C = E_n C = (BA)C = B(AC) = BE_n = B$. $\square$

1.2.11 Beispiel: (a) E_n ist invertierbar.

(b) Sei $A = \begin{pmatrix} 1 & 2 \\ 2 & 5 \end{pmatrix}$. A ist invertierbar, denn

$$\begin{pmatrix} 1 & 2 \\ 2 & 5 \end{pmatrix} \cdot \begin{pmatrix} 5 & -2 \\ -2 & 5 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 5 & -2 \\ -2 & 5 \end{pmatrix} \cdot \begin{pmatrix} 1 & 2 \\ 2 & 5 \end{pmatrix}.$$

(c) Allgemeiner gilt: Ist $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in M(2 \times 2, K)$ mit $ad - bc \neq 0$. Dann ist A invertierbar und

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

1.2.12 Lemma: (a) Das Inverse einer invertierbaren Matrix A ist invertierbar und es gilt $(A^{-1})^{-1} = A$.

(b) Das Produkt AB zweier invertierbarer $n \times n$-Matrizen A und B ist invertierbar und es gilt $(AB)^{-1} = B^{-1}A^{-1}$.

1.2.13 Wie invertiert man eine Matrix? Es gibt einen Algorithmus, der es einem erlaubt zu entscheiden, ob eine gegebene Matrix $A \in M(n \times n, K)$ invertierbar ist und der gegebenenfalls ihr Inverses berechnet.

Wir wollen ein $B = (b_{ij}) \in M(n \times n, K)$ finden mit $AB = E_n$. Betrachten wir nur die j-te Spalte so müssen wir also das lineare Gleichungssystem

$$(*) \qquad A \begin{pmatrix} b_{1j} \\ \vdots \\ b_{nj} \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix} \leftarrow j\text{-te Stelle}$$

in den Unbekannten $(b_{1j}, \ldots, b_{nj})$ lösen. Dazu wenden wir Gaußschen Algorithmus simultan auf alle Gleichungssysteme $(*)$ für $j = 1, \ldots, n$ an.

Wir führen den Algorithmus zuerst an zwei Beispielen vor.

1.2.14 Beispiel: (a) Sei $A = \begin{pmatrix} 0 & 1 & -4 \\ 1 & 2 & -1 \\ 1 & 1 & 2 \end{pmatrix}$. Wir versuchen A mit Hilfe von Zeilenumformungen in die Einheitsmatrix umzuwandeln und führen alle Umformungen parallel auch an E_3 durch.

$$\begin{array}{ccc|ccc} 0 & 1 & -4 & 1 & 0 & 0 \\ 1 & 2 & -1 & 0 & 1 & 0 \\ 1 & 1 & 2 & 0 & 0 & 1 \end{array}$$

$$\begin{array}{ccc|ccc} 1 & 2 & -1 & 0 & 1 & 0 \\ 0 & 1 & -4 & 1 & 0 & 0 \\ 1 & 1 & 2 & 0 & 0 & 1 \end{array}$$

$$\begin{array}{ccc|ccc} 1 & 2 & -1 & 0 & 1 & 0 \\ 0 & 1 & -4 & 1 & 0 & 0 \\ 0 & -1 & 3 & 0 & -1 & 1 \end{array}$$

$$\begin{array}{ccc|ccc} 1 & 2 & -1 & 0 & 1 & 0 \\ 0 & 1 & -4 & 1 & 0 & 0 \\ 0 & 0 & -1 & 1 & -1 & 1 \end{array}$$

$$\begin{array}{ccc|ccc} 1 & 2 & -1 & 0 & 1 & 0 \\ 0 & 1 & -4 & 1 & 0 & 0 \\ 0 & 0 & 1 & -1 & 1 & -1 \end{array}$$

$$\begin{array}{ccc|ccc} 1 & 0 & 7 & -2 & 1 & 0 \\ 0 & 1 & -4 & 1 & 0 & 0 \\ 0 & 0 & 1 & -1 & 1 & -1 \end{array}$$

$$\begin{array}{ccc|ccc} 1 & 0 & 0 & 5 & -6 & 7 \\ 0 & 1 & -4 & 1 & 0 & 0 \\ 0 & 0 & 1 & -1 & 1 & -1 \end{array}$$

$$\begin{array}{ccc|ccc} 1 & 0 & 0 & 5 & -6 & 7 \\ 0 & 1 & 0 & -3 & 4 & -4 \\ 0 & 0 & 1 & -1 & 1 & -1 \end{array}$$

$\Rightarrow A$ ist invertierbar mit Inversem $A^{-1} = \begin{pmatrix} 5 & -6 & 7 \\ -3 & 4 & -4 \\ -1 & 1 & -1 \end{pmatrix}$.

(b)

$$\begin{array}{ccc|ccc} 1 & 0 & 1 & 1 & 0 & 0 \\ 0 & -1 & 0 & 0 & 1 & 0 \\ 1 & 1 & 1 & 0 & 0 & 1 \end{array}$$

$$\begin{array}{ccc|ccc} 1 & 0 & 1 & 1 & 0 & 0 \\ 0 & -1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & -1 & 0 & 1 \end{array}$$

$$\begin{array}{ccc|ccc} 1 & 0 & 1 & 1 & 0 & 0 \\ 0 & -1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 & 1 & 1 \end{array}$$

$$\Rightarrow B\colon = \begin{pmatrix} 1 & 0 & 1 \\ 0 & -1 & 0 \\ 1 & 1 & 1 \end{pmatrix} \text{ ist nicht invertierbar.}$$

Wir haben also das folgende Verfahren zur Bestimmung des Inversen von $A \in M(n \times n, K)$: Man schreibe die Matrizen A und E_n nebeneinander. Alle Zeilenumformungen, die an A vorgenommen werden, führt man parallel auch an E_n durch. Zunächst bringt man A durch Zeilenumformungen auf Zeilenstufenform. Dabei stellt sich heraus ob $r = n$ ist. Ist $r < n$, so ist A nicht invertierbar. Ist $r = n$, so führt man weitere Zeilenumformungen durch bis aus A links die Matrix E_n geworden ist. Rechts steht dann die inverse Matrix A^{-1}. Das möchte ich jetzt begründen.

Die Zeilenumformungen, die wir in jedem Schritt vorgenommen haben waren jeweils von einem der folgenden Typen:

Typ 1: Vertauschung zweier Zeilen;

Typ 2: Multiplikation einer Zeile mit einem Zahlenfaktor $\neq 0$;

Typ 3: Addition des Vielfachen einer Zeile zu einer anderen.

Diese heißen elementare Zeilenumformungen. Analog definiert man elementare Spaltenumformungen des Typs 1, 2 und 3.

Eine $m \times m$ Matrix $\widetilde{E}$ heißt Elementarmatrix vom Typ i wenn sie aus der Einheitsmatrix E_m durch eine elementare Zeilenumformung vom Typ i hervorgeht. Die Elementarmatrizen haben die folgende Gestalt:

$$\begin{array}{cl} \left(\begin{array}{ccccccccccc} 1 & & & & & & & & & & 0 \\ & \ddots & & & & & & & & & \\ & & 1 & & & & & & & & \\ & & & 0 & & & & 1 & & & \\ & & & & 1 & & & & & & \\ & & & & & \ddots & & & & & \\ & & & & & & 1 & & & & \\ & & & 1 & & & & 0 & & & \\ & & & & & & & & 1 & & \\ & & & & & & & & & \ddots & \\ 0 & & & & & & & & & & 1 \end{array}\right) & \begin{array}{l} \\ \\ \\ i\text{-te Zeile} \\ \\ \\ \\ j\text{-te Zeile} \\ \\ \\ \\ \end{array} \end{array} \qquad \qquad \text{Typ 1}$$

$$\begin{pmatrix} 1 & & & & & & 0 \\ & \ddots & & & & & \\ & & 1 & & & & \\ & & & \lambda & & & \\ & & & & 1 & & \\ & & & & & \ddots & \\ 0 & & & & & & 1 \end{pmatrix} \qquad \text{Typ 2}$$

$$\begin{array}{c l} \left(\begin{matrix} 1 & & & & & & 0 \\ & \ddots & & & & & \\ & & 1 & \dots & 0 & & \\ & & \vdots & \ddots & \vdots & & \\ & & \lambda & \dots & 1 & & \\ & & & & & \ddots & \\ 0 & & & & & & 1 \end{matrix}\right) & \begin{matrix} \\ \\ i\text{-te Zeile} \\ \\ j\text{-te Zeile} \\ \\ \\ \end{matrix} \end{array} \qquad \text{Typ 3}$$

Sei $E_{ij} = E_{ij}^{(m)} \in M(m \times m, K)$ die Matrix, deren Eintrag an der Stelle (i,j) gleich 1 ist, und deren andere Einträge alle Null sind. Seien $i, j \in \{1, 2, \dots, m\}$ und sei $\lambda \in K$. Dann gilt:

(i) $\widetilde{E} = E_m + E_{ij} + E_{ji} - E_{ii} - E_{jj}$ ist die Elementarmatrix, die aus E_m durch Vertauschung der i-ten und j-ten Zeile hervorgeht.

(ii) $\widetilde{E} = E_m + (\lambda - 1)E_{ii}$ ist die Elementarmatrix, die aus E_m durch Multiplikation der i-ten Zeile mit $\lambda \in K$ hervorgeht.

(iii) $\widetilde{E} = E_m + \lambda E_{ji}$ ist die Elementarmatrix, die aus E_m durch Addition des λ-fachen der i-ten Zeile zur j-ten Zeile hervorgeht.

1.2.15 Satz: Entsteht $\widetilde{A} \in M(m \times n, K)$ aus $A \in M(m \times n, K)$ durch eine elementare Zeilenumformung und ist $\widetilde{E} \in M(m \times m, K)$ die Elementarmatrix, die aus E_m durch die selbe elementare Zeilenumformung entsteht, so gilt:

$$\widetilde{A} = \widetilde{E} A.$$

1.2.16 Beispiel: Sei $A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$. Betrachte die elementare Zeilenumformung “Addiere das Doppelte der ersten Zeile zur zweiten”. Also ist $\widetilde{A} \begin{pmatrix} 1 & 2 \\ 5 & 8 \end{pmatrix}$

und $\widetilde{E} = \begin{pmatrix} 1 & 0 \\ 2 & 1 \end{pmatrix}$ und es gilt:

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

Beweis von 1.2.15: Wir betrachten nur den Fall einer elementare Zeilenumformung vom Typ 3. Seien $z_1, \ldots, z_m \in K^n$ die Zeilen von A. $\widetilde{A}$ entstehe aus A durch Addition des λ-fachen der i-ten Zeile zur j-ten:

$$\Rightarrow \qquad \widetilde{A} = \begin{pmatrix} z_1 \\ \vdots \\ z_i \\ \vdots \\ z_j + \lambda z_i \\ \vdots \\ z_m \end{pmatrix} = \begin{pmatrix} e_1 A \\ \vdots \\ e_i A \\ \vdots \\ (e_j + \lambda e_i) A \\ \vdots \\ e_m A \end{pmatrix} = \begin{pmatrix} e_1 \\ \vdots \\ e_j + \lambda e_i \\ \vdots \\ e_m \end{pmatrix} A = \widetilde{E} A.$$

□

1.2.17 Folgerung: Die Elementarmatrizen sind invertierbar, die Inversen sind ebenfalls Elementarmatrizen.

Beweis (wieder nur im Fall einer elementare Zeilenumformung vom Typ 3): Sei $\widetilde{E_1}$ (bzw. $\widetilde{E_2}$) die $n \times n$-Matrix, die aus E_n durch Addition des λ-fachen bzw. des $-\lambda$-fachen) der i-ten Zeile zur j-ten hervorgeht. Nach Satz 1.2.15 ist das Produkt $\widetilde{E_2}\widetilde{E_1}$, die Matrix die aus E_n hervorgeht, indem man erst das λ-fache der i-ten Zeile zur j-ten addiert und dann nochmal das $-\lambda$-fache der i-ten Zeile zur j-ten addiert, d.h. es gilt $\widetilde{E_2}\widetilde{E_1} = E_n$. Ebenso sieht man dass $\widetilde{E_1}\widetilde{E_2} = E_n$. Also ist $\widetilde{E_1}$ invertierbar und es gilt $\widetilde{E_1}^{1} = \widetilde{E_2}$. □

1.2.18 Satz: Sei $A \in M(n \times n, K)$. Die folgenden Bedingungen sind äquivalent:

(i) A lässt sich durch elementare Zeilenumformungen in die Einheitsmatrix überführen.

(ii) A lässt sich als Produkt von Elementarmatrizen schreiben.

(iii) A ist invertierbar.

(iv) Das homogene lineare Gleichungssystem $A \cdot x = 0$ besitzt nur die triviale Lösung $x = 0$.

Beweis: (i) $\Rightarrow$ (ii) Läßt sich A durch Zeilenumformungen in die Einheitsmatrix überführen, so gibt es nach Satz 1.2.15 Elementarmatrizen $\widetilde{E}_1, \dots, \widetilde{E}_s$ mit

$$\widetilde{E}_s \cdot \ldots \cdot \widetilde{E}_1 \cdot A = E_n.$$

Nach 1.2.17 sind $\widetilde{E}_1, \dots, \widetilde{E}_s$ invertierbar und ihre Inversen $\widetilde{E}_1^{-1}, \dots, \widetilde{E}_s^{-1}$ ebenfalls Elementarmatrizen. Multiplizieren wir obige Gleichung von links mit $\widetilde{E}_1^{-1} \cdot \ldots \cdot \widetilde{E}_s^{-1}$ so erhalten wir

$$A = (\widetilde{E}_1^{-1} \cdot \ldots \cdot \widetilde{E}_s^{-1})(\widetilde{E}_s \cdot \ldots \cdot \widetilde{E}_1) \cdot A = \widetilde{E}_1^{-1} \cdot \ldots \cdot \widetilde{E}_s^{-1}.$$

d.h. A ist Produkt von Elementarmatrizen.

(ii) $\Rightarrow$ (iii) Folgt aus Lemma 1.2.12 (b) und 1.2.17.

(iii) $\Rightarrow$ (iv) Sei x eine Lösung des Gleichungssystems

$$A \cdot \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}. \tag{4}$$

Multiplizier (4) von links mit A^{-1}:

$$x = E_n \cdot x = (A^{-1}A)x = A^{-1}(Ax) = A^{-1}0 = 0$$

(iv) $\Rightarrow$ (i) Wir nehmen an, dass (4) nur trivial lösbar ist und verwandeln die Matrix A durch elementare Zeilenumformungen in eine Matrix $\widetilde{A} = (\widetilde{a}_{ij}$ in Zeilenstufenform. Da sich die Lösungsmenge eines linearen Gleichungssystems bei elementaren Zeilenumformungen nicht verändert, hat auch $\widetilde{A}x = 0$ nur die triviale Lösung $x = 0$. Nach 1.1.5 gibt es in $\widetilde{A}$ keine Nullzeile. Also ist $r = n$ und $\widetilde{A}$ ist eine *obere Dreiecksmatrix* deren Diagonaleinträge alle $\neq 0$ sind, d.h. $\widetilde{a}_{ij} = 0$ falls $i > j$ und $\widetilde{a}_{ii} \neq 0$ für $i = 1, 2, \dots, n$. Durch weitere elementare Zeilenumformungen vom Typ 2 läßt sich $\widetilde{A}$ in eine obere Dreiecksmatrix deren Diagonaleinträge alle $= 1$ sind überführen. Diese läßt sich schliesslich durch Zeilenumformungen vom Typ 3 in die Einheitsmatrix umwandeln. $\square$

1.2.19 Bemerkung: Sei $A \in M(n \times n, K)$ invertierbar. Nach obigem Satz 1.2.18 können wir A durch elementare Zeilenumformungen in die Einheitsmatrix überführen. Führt man dieselben elementare Zeilenumformungen parallel an der Einheitsmatrix durch so erhält man A^{-1}.

Beweis: Nach 1.2.15 entspricht ja jeder Umformungschritt der Multiplikation von links mit der entsprechenden Elementarmatrix. Gibt es insgesammt s Umformungen und ist $\widetilde{E}_i$ die Elementarmatrix, die zur i-ten Umformung gehört, so gilt:

$$\widetilde{E}_s \cdot \ldots \cdot \widetilde{E}_1 \cdot A = E_n$$

Durch dieselben Umformungen wird E_n in die Matrix $B\colon = \widetilde{E}_s \cdot \ldots \cdot \widetilde{E}_1 \cdot E_n = \widetilde{E}_s \cdot \ldots \cdot \widetilde{E}_1$ überführt. Es folgt:

$$B = \widetilde{E}_s \ldots \widetilde{E}_1 E_n = (\widetilde{E}_s \ldots \widetilde{E}_1)(AA^{-1}) = (\widetilde{E}_s \ldots \widetilde{E}_1 A)A^{-1} = E_n A^{-1} = A^{-1}$$

wie behauptet. □

1.2.20 Die Transponierte einer Matrix. Jeder $m \times n$ Matrix $A = (a_{ij})$ ist die an der Diagonalen $i = j$ gespiegelte Matrix A^t (die *Transponierte von* A) zugeordnet, deren i-te Zeile aus den Koeffizienten der i-ten Spalte von A besteht ($1 \le i \le n$). Explizit:

$$A = \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & & a_{2n} \\ \vdots & & & \vdots \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{pmatrix} \quad \Rightarrow \quad A^t = \begin{pmatrix} a_{11} & a_{21} & \ldots & a_{m1} \\ a_{12} & a_{22} & & a_{m2} \\ \vdots & & & \vdots \\ a_{1n} & a_{2n} & \ldots & a_{mn} \end{pmatrix} \in M(n \times m, K)$$

(d.h. $A^t = (b_{ij})_{\substack{i=1,\ldots,n \\ j=1,\ldots,m}}$ mit $b_{ij} = a_{ji}$).

1.2.21 Beispiel: $A = \begin{pmatrix} 1 & 0 & 3 \\ 2 & 5 & -1 \end{pmatrix} \Rightarrow A^t = \begin{pmatrix} 1 & 2 \\ 0 & 5 \\ 3 & -1 \end{pmatrix}$.

Es gelten die folgenden Rechenregeln

(1) $(A + B)^t = A^t + B^t \qquad \forall\, A, B \in M(m \times n, K)$;

(2) $(\alpha A)^t = \alpha A^t \qquad \forall\, \alpha \in K,\ A, \in M(m \times n, K)$;

(3) $(A^t)^t = A \qquad \forall\, A, \in M(m \times n, K)$;

(4) $(AB)^t = B^t A^t \qquad \forall\, A \in M(m \times n, K),\ B \in M(n \times r, K)$.

(5) $A \in M(n \times n, K)$ ist genau dann invertierbar wenn A^t invertierbar ist und es gilt $(A^t)^{-1} = (A^{-1})^t$.

1.2.22 Folgerung: Entsteht $\widetilde{A} \in M(m \times n, K)$ aus $A \in M(m \times n, K)$ durch eine elementare Spaltenumformung und ist $\widetilde{E} \in M(m \times m, K)$ die Elementarmatrix, die aus E_n durch die selbe elementare Spaltenumformung entsteht, so gilt:

$$\widetilde{A} = A\widetilde{E}.$$

Beweis: Nach 1.2.15 gilt für die Matrix $\widetilde{A}^t$:

$$\widetilde{A}^t = \widetilde{E}^t A^t = (A\widetilde{E})^t$$

also $\widetilde{A} = A\widetilde{E}$. □

1.3 Determinanten

Einer quadratischen Matrix kann man eine Zahl zuordnen – ihre Determinante – die genau dann nicht Null ist, wenn die Matrix invertierbar ist. Man kann mit Hilfe von Determinanten die Lösung $x_1, \ldots, x_n$ eines linearen Gleichungssystems der Form

$$\begin{pmatrix} a_{11} & \ldots & a_{1n} \\ a_{21} & & a_{2n} \\ \vdots & & \vdots \\ a_{n1} & \ldots & a_{nn} \end{pmatrix} \begin{pmatrix} x_1 & \\ x_2 & \\ \vdots & \vdots \\ x_n & \end{pmatrix} := \begin{pmatrix} b_1 & \\ b_2 & \\ \vdots & \vdots \\ b_n & \end{pmatrix}$$

direkt als Funktion der Koeffizienten a_{ij}, b_i darstellen. Im Falle einer Matrix mit reellen Komponenten misst die Determinante das Volumen des von den Spalten (oder von Zeilen) erzeugten *Parallepiped* (ein 2-dimensionales Parallepiped ist ein Parallelogramm, ein 3-dimensionales ein Spat).

1.3.1 2-reihige Determinanten Sei $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ eine 2×2-Matrix. Die Zahl

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

heißt Determinante von A.

Wir wollen die Nützlichkeit dieses Begriffs an den oben erwähnten Beispielen – lineare Gleichungssysteme und Flächeninhalt eines Parallelogramms – dokumentieren.

1.3.2 Sei

$$(*) \qquad \begin{array}{ccccc} a_{11}x_1 & + & a_{12}x_2 & = & b_1 \\ a_{21}x_1 & + & a_{22}x_2 & = & b_2 \end{array}$$

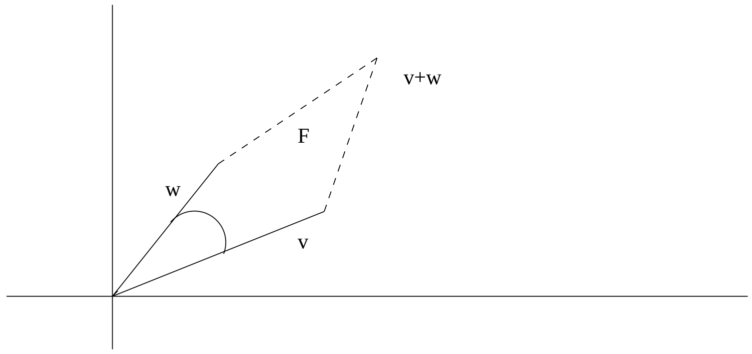
ein lineares Gleichungssystem von dem wir annehmen, dass die Koeffizientenmatrix $A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$ invertierbar ist. Also ist $\det A = a_{11}a_{22} - a_{12}a_{21} \neq 0$ und das Inverse ist $A^{-1} = \frac{1}{\det A}\begin{pmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{pmatrix}$.

Wir suchen eine allgemeine Formel für die Lösung. Multiplizieren wir die Matrizengleichung $A\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}$ mit A^{-1}, so erhalten wir für die Lösung von $(*)$:

$$x_1 = \frac{a_{22}b_1 - a_{12}b_2}{\det(A)} = \frac{\det\begin{pmatrix} b_1 & a_{12} \\ b_2 & a_{22} \end{pmatrix}}{\det\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}}$$

$$x_2 = \frac{-a_{21}b_1 + a_{11}b_2}{\det(A)} = \frac{\det\begin{pmatrix} a_{11} & b_1 \\ a_{21} & b_2 \end{pmatrix}}{\det\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}}$$

1.3.3 Seien $v = (a, b), w = (c, d) \in \mathbb{R}^2$. Wir wollen den Flächeninhalt F des von v und w aufgespannten Parallelogramms $\{sv + tw \mid 0 \leq s, t \leq 1\}$ bestimmen. Im obigen Bild bezeichnet φ den Winkel um den man v drehen



muss (entgegen dem Uhrzeigersinn), um einen zu w parallelen Vektor zu erhalten. Der Flächeninhalt ist Grundfläche mal Höhe. Schreiben wir v und w in Polarkoordinaten

$$\begin{aligned} v &= r_1\ (\cos\alpha, \sin\alpha) \\ w &= r_2\ (\cos\beta, \sin\beta) \end{aligned}$$

so gilt also

$$F = r_1 h = r_1 r_2 \sin\varphi$$

Dabei ist h der Abstand von w von der Geraden durch v und den Ursprung, also $h = r_2 \sin\varphi = r_2 \sin(\beta - \alpha)$.

$$\Rightarrow \qquad F = hr = r_1 r_2 (\sin\beta\cos\alpha - \sin\alpha\cos\beta) = \det\begin{pmatrix} r_1\cos\alpha & r_1\sin\alpha \\ r_2\cos\beta & r_2\sin\beta \end{pmatrix} = \det\begin{pmatrix} v \\ w \end{pmatrix}.$$

1.3.4 Sei $K = \mathbb{Q}$ oder $\mathbb{R}$. Wir listen jetzt einige Eigenschaften der Abbildung

$$\det : M(2 \times 2, K) \longrightarrow K, A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \mapsto ad - bc = \det A$$

auf. Dabei benutzen wir die Notation $A = \begin{pmatrix} v \\ w \end{pmatrix}$ wobei $v = (a, b), w = (c, d)$ die Zeilenvektoren von V sind.

(a) Für $\lambda, \mu \in K$ ist

$$\det\begin{pmatrix} \lambda v \\ w \end{pmatrix} = \lambda \cdot \det\begin{pmatrix} v \\ w \end{pmatrix}, \quad \det\begin{pmatrix} v \\ \mu w \end{pmatrix} = \mu \cdot \det\begin{pmatrix} v \\ w \end{pmatrix}.$$

Im Falle $K = \mathbb{R}$ folgt das aus der Tatsache, dass der Fläche eines Parallelogramms so gestreckt wird wie die einzelnen Seiten.

(b) Für $w_1, w_2, v \in K^2$ ist

$$\det\begin{pmatrix} v \\ w_1 + w_2 \end{pmatrix} = \det\begin{pmatrix} v \\ w_1 \end{pmatrix} + \det\begin{pmatrix} v \\ w_2 \end{pmatrix}$$

Ebenso gilt für $v_1, v_2, w \in K^2$

$$\det\begin{pmatrix} v_1 + v_2 \\ w \end{pmatrix} = \det\begin{pmatrix} v_1 \\ w \end{pmatrix} + \det\begin{pmatrix} v_2 \\ w \end{pmatrix}.$$

(c)

$$\det\begin{pmatrix} e_1 \\ e_2 \end{pmatrix} = 1.$$

Das bedeutet (für $K = \mathbb{R}$), dass das Einheitsquadrat den Flächeninhalt 1 hat. (d)

$$\det \begin{pmatrix} v \\ v \end{pmatrix} = 0.$$

(e)

$$\det \begin{pmatrix} v \\ w \end{pmatrix} = -\det \begin{pmatrix} w \\ v \end{pmatrix}.$$

(f) $\quad \det \begin{pmatrix} v \\ w \end{pmatrix} \neq 0 \iff \begin{pmatrix} v \\ w \end{pmatrix}$ ist invertierbar.

1.3.5 Bevor wir Determinanten von 3×3-Matrizen einführen soll noch einmal an die aus der Schulmathematik bekannten Skalar-, Vektor- und Spatprodukte im $\mathbb{R}^3$ erinnert werden. Wir fassen dabei Vektoren im $\mathbb{R}^3$ vorübergehend als Spaltenvektoren auf.

Das Skalarprodukt $a \bullet b$ zweier Vektoren $a = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix}, b = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix}$ ist definiert durch

$$a \bullet b = a_1b_1 + a_2b_2 + a_3b_3.$$

Sind weder a noch b gleich dem Nullvektor und ist $\varphi \in [0, \pi]$ der Winkel zwischen a und b, so gilt

$$a \bullet b = \|a\| \cdot \|b\| \cdot \cos(\varphi).$$

Dabei bezeichnet $\|a\|$ (bzw. $\|b\|$) die Länge des Vektors a (bzw. b). Es gilt also $\|a\| = \sqrt{a_1^2 + a_2^2 + a_3^2} = \sqrt{a \bullet a}$.

Das Vektorprodukt $a \times b$ ist definiert durch

$$a \times b = \begin{pmatrix} a_2b_3 - a_3b_2 \\ a_3b_1 - a_1b_3 \\ a_1b_2 - a_2b_1 \end{pmatrix}.$$

Der Vektor $a \times b \in \mathbb{R}^3$ ist der Nullvektor, falls a und b linear abhängig sind. Andernfalls ist er durch die folgenden Eigenschaften eindeutig charakterisiert:

- $a \times b$ steht senkrecht auf a und b;
- $a, b, a \times b$ bildet ein Rechtssystem;

- Die Länge von $a \times b$ ist gleich dem Flächinhalt des von a und b aufgespannten Parallelogramms.

Das Spatprodukt von drei Vektoren $a = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix}, b = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix}$ und $c = \begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix}$ ist gegeben durch

$$[a, b, c] \colon = a \bullet (b \times c).$$

Bis auf das Vorzeichen, ist es gleich dem Volumen des von a, b und c aufgespannten Spats $\{t_1 a + t_2 a + t_3 a \mid 0 \leq t_1, t_2, t_3 \leq 1\}$. Das sieht man wie folgt. Der Flächeninhalt der Grundfläche des Spats mit den Kanten b und c ist $F = \|b \times c\|$. Der Vektor $v \colon = \frac{1}{F} \cdot b \times c$ hat die Länge 1 und steht senkrecht auf dieser Grundfläche. Die Höhe h des Spats bzgl. dieser Grundfläche ist $h = |a \bullet v|$. Damit ergibt sich

$$V = hF = |Fa \bullet v| = |a \bullet (b \times c)|.$$

1.3.6 3-reihige Determinanten

Sei

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$$

eine relle 3×3 Matrix. Wir bezeichnen die 3 Spaltenvektoren von A mit

$$a_1 = \begin{pmatrix} a_{11} \\ a_{21} \\ a_{31} \end{pmatrix}, \quad a_2 = \begin{pmatrix} a_{12} \\ a_{22} \\ a_{32} \end{pmatrix}, \quad a_3 = \begin{pmatrix} a_{13} \\ a_{23} \\ a_{33} \end{pmatrix}$$

Die Determinante von A ist das Spatprodukt $[a_1, a_2, a_3]$ also

$$\begin{aligned} \det A :=& a_1 \bullet (a_2 \times a_3) = a_1 \bullet \begin{pmatrix} a_{22}a_{33} & -a_{23}a_{32} \\ -a_{12}a_{33} & +a_{13}a_{32} \\ a_{12}a_{23} & -a_{13}a_{22} \end{pmatrix} \\ =& a_{11}(a_{21}a_{33} - a_{23}a_{32}) - a_{21}(a_{12}a_{33} - a_{13}a_{32}) + a_{31}(a_{12}a_{23} - a_{13}a_{22}) \\ =& a_{11} \det \begin{pmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{pmatrix} - a_{21} \det \begin{pmatrix} a_{12} & a_{13} \\ a_{32} & a_{33} \end{pmatrix} + a_{31} \det \begin{pmatrix} a_{12} & a_{13} \\ a_{22} & a_{23} \end{pmatrix} \end{aligned}$$

Bis auf das Vorzeichen ist $\det A$ gleich dem Volumen des von des Spalten aufgespannten Spats.

Berechnung der Determinante einer 3×3 Matrix mit Hilfe der *Formel von Sarrus*:

$$\begin{array}{ccccc} + & + & + & & \\ a_{11} & a_{12} & a_{13} & a_{11} & a_{12} \\ a_{21} & a_{22} & a_{23} & a_{21} & a_{22} \\ a_{31} & a_{32} & a_{33} & a_{31} & a_{32} \\ - & - & - & & \end{array}$$

Man schreibt die ersten beiden Spalten noch einmal rechts neben A, addiert längs der $\searrow$-Diagonalen zu bildenden Produkte und subtrahiert die längs der $\nearrow$ berechneten Produkte. Also:

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

1.3.7 Beispiel: $A = \begin{pmatrix} 0 & -2 & 3 \\ -2 & 1 & -2 \\ 3 & 6 & 5 \end{pmatrix}$

$$\begin{array}{ccccc} 0 & -2 & 3 & 0 & -2 \\ -2 & 1 & -2 & -2 & 1 \\ 3 & 6 & 5 & 3 & 6 \end{array}$$

$$\det A = 12 - 36 - 9 - 20 = -53.$$

Wir betrachten jetzt $n \times n$ Matrizen über einem beliebigen Körper K. Ist $A \in M(n \times n, K$ mit den Zeilenvektoren $a_1, \ldots, a_n$ so schreiben wir

$$A = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}.$$

1.3.8 Satz und Definition: Es gibt genau eine Abbildung

$$\det : M(n \times n, K) \to K$$

mit den folgenden Eigenschaften:

(D1) Für jeden Zeilenindex $i \in \{1, \ldots, n\}$ gilt:

(a) Ist $a_i = a'_i + a''_i$ so ist

$$\det \begin{pmatrix} a_1 \\ \vdots \\ a_i \\ \vdots \\ a_n \end{pmatrix} = \det \begin{pmatrix} a_1 \\ \vdots \\ a'_i \\ \vdots \\ a_n \end{pmatrix} + \det \begin{pmatrix} a_1 \\ \vdots \\ a''_i \\ \vdots \\ a_n \end{pmatrix}.$$

(b) Ist $a_i = \lambda a'_i, \lambda \in K$ so ist

$$\det \begin{pmatrix} a_1 \\ \vdots \\ a_i \\ \vdots \\ a_n \end{pmatrix} = \lambda \det \begin{pmatrix} a_1 \\ \vdots \\ a'_i \\ \vdots \\ a_n \end{pmatrix}.$$

(D2) Hat A zwei gleiche Zeilen, so ist

$$\det A = 0.$$

(D3) $\det E_n = 1$.

det heißt *Determinante*, $\det A$ die Determinante von A.

Die Eigenschaft (D1) bedeutet daß det *linear* in den Zeilen ist. Dieser Begriff wird im nächsten Abschnitt eingeführt und untersucht.

Bevor wir zum Beweis des Satzes kommen sollen einige Folgerungen aus den Axiomen (D1) – (D3) abgeleitet werden.

1.3.9 Lemma (Änderung von det bei elementaren Zeilenumformungen): Sei $\det : M(n \times n, K) \to K$ eine Abbildung mit den Eigenschaften (D1) und (D2). Seien $A, \widetilde{A} \in M(n \times n, K)$.

(a) Entsteht $\widetilde{A}$ aus A durch Vertauschung zweier Zeilen so ist $\det \widetilde{A} = -\det A$.

(b) Entsteht $\widetilde{A}$ aus A durch Multiplikation einer Zeile mit $\lambda \in K$, so gilt $\det \widetilde{A} = \lambda \det A$.

(c) Entsteht $\widetilde{A}$ aus A durch Addition des Vielfachen einer Zeile zu einer anderen so gilt: $\det \widetilde{A} = \det A$.

Beweis: (b) folgt direkt aus (D1).

zu (c): Entsteht $\widetilde{A}$ aus $A = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}$ durch Addition des λ-fachen der i-ten Zeile zur j-ten, so folgt aus der Eigenschaft (D1) für die j-te Zeile:

$$\det \widetilde{A} = \det \begin{pmatrix} \vdots \\ a_i \\ \vdots \\ \lambda a_i + a_j \\ \vdots \end{pmatrix} = \det \begin{pmatrix} \vdots \\ a_i \\ \vdots \\ \lambda a_i \\ \vdots \end{pmatrix} + \det \begin{pmatrix} \vdots \\ a_i \\ \vdots \\ a_j \\ \vdots \end{pmatrix}$$

$$= \lambda \det \begin{pmatrix} \vdots \\ a_i \\ \vdots \\ a_i \\ \vdots \end{pmatrix} + \det A \overset{(D2)}{=} \det A.$$

zu (a): Seien die i-te und die j-te die beiden zu vertauschenden Zeilen. Nach (b) und (c) ergibt sich durch mehrmaliges Anwenden von Zeilenumformungen:

$$\det A = \det \begin{pmatrix} \vdots \\ a_i \\ \vdots \\ a_j \\ \vdots \end{pmatrix} = \det \begin{pmatrix} \vdots \\ a_i \\ \vdots \\ a_j + a_i \\ \vdots \end{pmatrix} = \det \begin{pmatrix} \vdots \\ a_i - (a_i + a_j) \\ \vdots \\ a_j + a_i \\ \vdots \end{pmatrix}$$

$$= -\det \begin{pmatrix} \vdots \\ a_j \\ \vdots \\ a_j + a_i \\ \vdots \end{pmatrix} = -\det \begin{pmatrix} \vdots \\ a_j \\ \vdots \\ a_i + a_j + (-1)a_j \\ \vdots \end{pmatrix} = -\det \widetilde{A}$$

$\square$

1.3.10 Folgerung: Sei $\det : M(n \times n, K) \to K$ eine Abbildung mit den Eigenschaften (D1) – (D3). Entsteht $\widetilde{A}$ aus $A \in M(n \times n, K)$ durch eine elementare Zeilenumformung und ist $\widetilde{E} \in M(n \times n, K)$ die zugehörige Elementarmatrix (also $\widetilde{A} = \widetilde{E}A$) so gilt:

$$\det \widetilde{A} = \det \widetilde{E} \det A$$

Beweis: Nach dem Lemma unterscheiden sich $\det \widetilde{A}$ und $\det A$ durch einen Zahlenfaktor α (d.h. $\det \widetilde{A} = \alpha \det A$) der nur vom Typ der Zeilenumformung und nicht von A selber abhängt. Also gilt auch $\det \widetilde{E} = \alpha \det E_n$ und daher wegen (D3) $\det \widetilde{E} = \alpha$.

$$\Rightarrow \qquad \det \widetilde{A} = \det \widetilde{E} \det A.$$

□

1.3.11 Folgerung: Ist $A \in M(n \times n, K)$ ein Produkt von Elementarmatrizen $A = \widetilde{E}_1 \cdot \ldots \cdot \widetilde{E}_s$ so gilt:

$$\det A = \det \widetilde{E}_1 \cdot \ldots \cdot \det \widetilde{E}_s.$$

1.3.12 Folgerung: Ist $\det : M(n \times n, K) \to K$ eine Abbildung mit den Eigenschaften (D1) – (D3) und ist $A \in M(n \times n, K)$ so gilt:

$$A \text{ ist invertierbar} \quad \Leftrightarrow \quad \det A \neq 0.$$

Beweis: Sei A zunächst invertierbar. Nach Satz 1.2.18 läßt sich A als Produkt von Elementarmatrizen darstellen:

$$\widetilde{E}_s \cdot \ldots \cdot \widetilde{E}_1 A$$

Es genügt also die Behauptung für Elementarmatrizen zu zeigen. Das folgt aber sofort aus 1.3.9 (angewandt auf die Einheitsmatrix).

Gelte nun umgekehrt $\det A \neq 0$. Durch elementare Zeilenumformungen vom Typ 1 und 3 läßt sich A in eine Matrix $\widetilde{A}$ in Zeilenstufenform umwandeln. Nach Lemma 1.3.9 gilt $\det \widetilde{A} = \pm \det A \neq 0$. Also enthält $\widetilde{A}$ keine Nullzeile und die “Stufen” haben alle die Breite 1, d.h. $\widetilde{A}$ ist eine obere Dreiecksmatrix, deren Diagonaleinträge alle $\neq 0$ sind. Folglich können wir $\widetilde{A}$ durch weitere elementare Zeilenumformungen vom Typ 2 und 3 können in die Einheitsmatrix umwandeln. Nach 1.2.18 ist A invertierbar. □

Beweis von Satz 1.3.8:

1. Beweis der Eindeutigkeit: Seien det und det$'$ zwei Abbildungen mit den Eigenschaften (D1) – (D3) und sei $A \in M(n \times n, K)$. Wir müssen zeigen, daß $\det A = \det' A$. Ist zunächst A nicht invertierbar, so gilt nach obiger Folgerung:

$$\det A = 0 = \det{}' A.$$

Ist andererseits A invertierbar und läßt sich also durch elementare Zeilenumformungen in die Einheitsmatrix überführen. Nach Lemma 1.3.9 ändern sich det und det$'$ bei jeder dieser Umformungen um den gleichen Zahlenfaktor. Daher folgt $\det A = \det' A$ aus

$$\det E_n = 1 = \det{}' E_n$$

2. Beweis der Existenz: Die Existenz einer Abbildung $\det : M(n \times n, K) \to K$ mit den Eigenschaften (D1) – (D3) wird durch Induktion bewiesen. Für $n = 1$ hat offenbar $\det(a) := a$ diese Eigenschaften. Für den Induktionsschluß $n-1 \to n$ nehmen wir an, daß die Determinante für $(n-1)\times(n-1)$-Matrizen bereits existiert.

Für $A = (a_{ij}) \in M(n \times n, K)$ bezeichne A_{ij} die aus A durch weglassen der i-ten Zeile und j-ten Spalte entstandene $(n-1) \times (n-1)$ Matrix.

Wir fixieren jetzt einen Spaltenindex $j \in \{1, \ldots, n\}$ und definieren

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

Es ist leicht zu sehen, dass die so definierte Abbildung $\det : M(n\times n, K) \to K$ die Eigenschaften (D1) – (D3) hat.

1.3.13 Satz (Rechenregeln für Determinanten):

(a) (Multiplikationssatz) $\det AB = \det A \det B$ für alle $A, B \in M(n \times n, K)$.

(b) $\det A = \det A^t$ für alle $A \in M(n \times n, K)$. Insbesondere ist $\det A$ auch linear in den Spalten.

(c) (Entwicklung von $\det A$ nach einer beliebigen Zeile oder Spalte): Für $A = (a_{ij}) \in M(n \times n, K)$ gilt:

$$\det A = \sum_{i=1}^{n} (-1)^{i+j} a_{ij} \det A_{ij} \qquad \text{(Entwicklung von } \det A \text{ nach der } j\text{-ten Spalte),}$$

$$\det A = \sum_{j=1}^{n} (-1)^{i+j} a_{ij} \det A_{ij} \qquad \text{(Entwicklung nach der } i\text{-ten Zeile).}$$

Zum Beweis benötigen wir folgendes

1.3.14 Lemma: Für eine quadratische Matrix $A \in M(n \times n, K)$ sind die folgenden Bedingungen äquivalent:

(i) A ist invertierbar.

(ii) Es gibt eine Matrix $B \in M(n \times n, K)$ mit $BA = E_n$ (d.h. es A besitzt eine *Linksinverse*).

(iii) Es gibt eine Matrix $C \in M(n \times n, K)$ mit $AC = E_n$ (d.h. C ist *Rechstinverse* von A).

Beweis: (i) $\Rightarrow$ (ii), (iii) klar.

(ii) $\Rightarrow$ (i) Der Spaltenvektor $x \in M(n \times 1, K)$ sei Lösung des linearen Gleichungssystem $A \cdot x = 0$. Indem wir die Gleichung von Links mit B multiplizieren erhalten wir:

$$x = E_n x = (BA)x = B(Ax) = B0 = 0$$

d.h. die Gleichung $A \cdot x = 0$ besitzt nur die triviale Lösung $x = 0$. Nach 1.2.18 ist A invertierbar.

(iii) $\Rightarrow$ (i) Aus $AC = E_n$ folgt $C^t A^t = (AC)^t = E_n^t = E_n$ folgt aufgrund der Äquivalenz von (i) und (ii), dass A^t invertierbar ist. Indem wir in der Gleichung $A^t(A^t)^{-1} = E_n = (A^t)^{-1}A^t$ wieder zu den Transponierten übergehen, erhalten wir:

$$((A^t)^{-1})^t A = ((A^t)^{-1})^t (A^t)^t = (A^t(A^t)^{-1})^t = E_n = A((A^t)^{-1})^t.$$

Also ist A invertierbar und es gilt $A^{-1} = ((A^t)^{-1})^t$, d.h. $(A^{-1})^t = ((A^t)^{-1})$.

Die letzte Aussage folgt aus Lemma 1.2.10. $\square$

Beweis von Satz 1.3.13: (a) Wir zeigen zunächst, dass das Produkt $\det A \det B$ genau dann $= 0$ ist, wenn $\det AB = 0$. Nach 1.3.12 genügt es dazu die Äquivalenz der folgenden beiden Aussagen zu zeigen:

(i) A und B sind ivertierbar.

(ii) AB ist invertierbar.

(i) $\Rightarrow$ (ii): gilt nach 1.2.12.

(ii) $\Rightarrow$ (i): Wir zeigen, dass A die Bedingung (iii) und B die Bedingung (ii) in 1.3.14 erfüllt. Mit $X: = B(AB)^{-1}$ und $Y: = (AB)^{-1}A$ gilt nämlich:

$$AX = (AB)(AB)^{-1} = E_n, \qquad YB = (AB)^{-1}(AB) = E_n.$$

Sei jetzt $\det A \neq 0 \neq \det B$, d.h. A und B sind invertierbar und lassen sich folglich jeweils als Produkt von Elementarmatrizen darstellen:

$$A = \widetilde{E}_1 \ldots \widetilde{E}_s, \quad B = \widetilde{E}_{s+1} \ldots \widetilde{E}_{s+t}$$

Mit 1.3.10, (b) ergibt sich

$$\det AB = \det \widetilde{E}_1 \cdot \ldots \cdot \det \widetilde{E}_{s+t} = \det A \det B$$

(b) Ist A (also auch A^t) nicht invertierbar , so gilt nach 1.3.11

$$\det A = 0 = \det A^t$$

Ist A dagegen invertierbar, so schreiben wir A wieder als Produkt von Elementarmatrizen

$$A = \widetilde{E}_1 \cdot \ldots \cdot \widetilde{E}_s, \quad (\text{ also } A^t = \widetilde{E}_s^t \ldots \widetilde{E}_1^t.)$$

Aufgrund von 1.3.11 genügt es (ii) daher für eine Elementarmatrix $\widetilde{E}$ zu zeigen:

$$\det \widetilde{E} = \det \widetilde{E}^t$$

Das ist klar falls $\widetilde{E}$ vom Typ 1 oder 2 ist, denn dann ist $\widetilde{E}$ symmetrisch. Ist $\widetilde{E}$ vom Typ 3 also von der Form

$$\widetilde{E} = \begin{array}{cl} \begin{pmatrix} 1 & & & & & & 0 \\ & \ddots & & & & & \\ & & 1 & \ldots & 0 & & \\ & & \vdots & \ddots & \vdots & & \\ & & \lambda & \ldots & 1 & & \\ & & & & & \ddots & \\ 0 & & & & & & 1 \end{pmatrix} & \begin{array}{l} \\ \\ i\text{-te Zeile} \\ \\ j\text{-te Zeile} \\ \\ \\ \end{array} \end{array}$$

so ist

$$\widetilde{E}^t = \begin{array}{cl} \begin{pmatrix} 1 & & & & & & 0 \\ & \ddots & & & & & \\ & & 1 & \ldots & \lambda & & \\ & & \vdots & \ddots & \vdots & & \\ & & 0 & \ldots & 1 & & \\ & & & & & \ddots & \\ 0 & & & & & & 1 \end{pmatrix} & \begin{array}{l} \\ \\ i\text{-te Zeile} \\ \\ j\text{-te Zeile} \\ \\ \\ \end{array} \end{array}$$

wieder eine Elementarmatrix vom Typ 3. Mit Lemma 1.3.9 (c) folgt:

$$\det \widetilde{E} = \det E_n = 1 = \det \widetilde{E}^t$$

(c) Die Formel für die Entwicklung von $\det A$ nach der j-ten Spalte haben wir bereits gezeigt. Die Formel für die Entwicklung nach der i-ten Zeile folgt aus $\det A = \det A^t$. □

1.3.15 Beispiel: $A = \begin{pmatrix} 5 & -4 & 2 \\ 0 & 1 & 0 \\ 7 & 8 & 3 \end{pmatrix}$. Wir entwickeln $\det A$ nach der zweiten Zeile:

$$\det A = \det \begin{pmatrix} 5 & 2 \\ 7 & 3 \end{pmatrix} = 1.$$

1.3.16 Bemerkungen: (a) Für eine invertierbare $n \times n$ Matrix A gilt:

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

(b) Ist $A = (a_{ij}) \in M(n \times n, K)$ eine *obere* (bzw. *untere*) *Dreiecksmatrix*, d.h. ist $a_{ij} = 0 \quad \forall\, i > j$ (bzw. $a_{ij} = 0 \quad \forall\, j > i$) so gilt:

$$\det A = a_{11} \cdot a_{22} \cdot \ldots \cdot a_{nn}$$

1.3.17 Die Adjunkte von A: Sei $A = (a_{ij}) \in M(n \times n, K)$. Setze für $i, j \in \{1, \ldots, n\}$

$$a^*_{ij} := (-1)^{i+j} \det A_{ji}.$$

Die Matrix $A^* = (a^*_{ij})$ heißt *Adjunkte* von A. Nach Satz 1.3.13, (c) gilt für den Eintrag $(AA^*)_{ii}$ an der Stelle (i, i) des Produkts AA^*:

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

Für $j \neq i$ ergibt sich für $(AA^*)_{ij}$:

$$\sum_{k=1}^{n} a_{ik} a^*_{kj} = \sum_{k=1}^{n} a_{ik} (-1)^{k+j} \det A_{jk}$$

Das ist die Determinante der Matrix $\widetilde{A}$ die man aus A erhält, wenn man die j-te Zeile durch die i-te ersetzt. Also hat $\widetilde{A}$ zwei gleiche Zeilen und $\sum\limits_{k=1}^{n} a_{ik} a^*_{kj} = 0$.

Für das Produkt AA^* gilt also

$$AA^* = \begin{pmatrix} \det A & & 0 \\ & \ddots & \\ 0 & & \det A \end{pmatrix} = \det(A) \cdot E_n$$

Analog zeigt man: $A^*A = \det(A) \cdot E_n$. Damit ergibt sich:

1.3.18 Satz: Sei $A \in M(n \times n, K)$ invertierbar. Dann gilt für die inverse Matrix A^{-1}:

$$A^{-1} = \det(A)^{-1} A^* = \left(\frac{(-1)^{i+j} \det A_{ji}}{\det A} \right)_{i,j}.$$

Man kann mit Hilfe des Determinantenkalküls eine Formel für die Lösung eines linearen Gleichungssystems mit n Gleichungen in n Unbestimmten und invertierbarer Koeffizientenmatrix angeben.

1.3.19 Satz (Cramersche Regel): Sei

$$(*) \qquad \begin{array}{ccccccc} a_{11}x_1 & + & \dots & + & a_{1n}x_n & = & b_1 \\ \vdots & & & & \vdots & & \vdots \\ a_{n1}x_1 & + & \dots & + & a_{nn}x_n & = & b_n \end{array}$$

ein lineares Gleichungssystem mit n Gleichungen in n Unbestimmten mit Koeffizientenmatrix $A = (a_{ij}) \in M(n \times n, K)$. Ist $\det A \neq 0$ so besitzt $(*)$ eine eindeutig bestimmte Lösung $(x_1, \dots, x_n)$. Es gilt

$$x_i = \frac{\det \begin{pmatrix} a_{11} & \dots & b_1 & \dots & a_{1n} \\ \vdots & & \vdots & & \vdots \\ a_{n1} & \dots & b_n & \dots & a_{nn} \end{pmatrix}}{\det \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \dots & a_{nn} \end{pmatrix}}$$

wobei die Matrix im Zähler aus A entsteht, indem man die i-te Spalte von A durch $b := \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}$ ersetzt.

Beweis: Dass $(*)$ eine eindeutig bestimmte Lösung $(x_1, \ldots, x_n)$ besitzt, folgt aus der Invertierbarkeit von A. Seien $a_1, \ldots, a_n$ die Spalten von A. Sei $\widetilde{A}$ die Matrix mit den Spalten $a_1, \ldots, a_{i-1}, x_i \cdot a_i - b, a_{i+1}, \ldots, a_n$. Das homogene lineare Gleichungssystem

$$\widetilde{A} \begin{pmatrix} \widetilde{x}_1 \\ \vdots \\ \widetilde{x}_n \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}$$

in den Unbekannten $\widetilde{x}_1, \ldots, \widetilde{x}_n$ besitzt die Lösung

$$\widetilde{x}_1 \colon = x_1, \ldots, \widetilde{x}_{i-1} \colon = x_{i-1}, \widetilde{x}_i \colon = 1, \widetilde{x}_{i+1} \colon = x_{i+1} \ldots, \widetilde{x}_n \colon = x_n.$$

Nach Satz 1.2.18 ist $\widetilde{A}$ nicht invertierbar. Es folgt

$$\begin{aligned} 0 &= \det(a_1, \ldots, a_{i-1}, x_i \cdot a_i - b, a_{i+1}, \ldots, a_n) \\ &= x_i \det A - \det(a_1, \ldots, a_{i-1}, b, a_{i+1}, \ldots, a_n) \end{aligned}$$

wobei wir für die zweite Gleichheit benutzt haben, dass die Determinante auch in den Spalten linear ist. □

1.3.20 Beispiel: Wir betrachten das Gleichungssystems

$$\begin{aligned} x_1 + x_2 &= 1 \\ x_2 + x_3 &= 1 \\ 3x_1 + 2x_2 + x_3 &= 0 \end{aligned}$$

Für die Determinante der Koeffizientenmatrix gilt

$$\det \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 3 & 2 & 1 \end{pmatrix} = 1 \det \begin{pmatrix} 1 & 1 \\ 2 & 1 \end{pmatrix} + 3 \det \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} = 2$$

Also ergibt sich:

$$\begin{aligned} x_1 &= \frac{1}{2} \det \begin{pmatrix} 1 & 1 & 0 \\ 1 & 1 & 1 \\ 0 & 2 & 1 \end{pmatrix} = \frac{-2}{2} = -1 \\ x_2 &= \frac{1}{2} \det \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 3 & 0 & 1 \end{pmatrix} = \frac{1+3}{2} = 2 \\ x_3 &= \frac{1}{2} \det \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 3 & 2 & 0 \end{pmatrix} = \frac{-2+0}{2} = -1 \end{aligned}$$

Für große n ist die Cramersche Regel unpraktisch, da der Rechenaufwand zur Berechnung von $n+1$ Determinanten von $n \times n$-Matrizen von zu gross ist.

Zum Abschluss des Kapitels über Determinanten wollen wir noch die sogenannte Leibnizformel herleiten, die die Determinante einer quadratischen Matrix vollständig durch ihre Einträge ausdrückt. Dazu müssen wir zunächst an den Begriff des Vorzeichens (oder *Signum*) einer Permutation einführen.

Vorzeichen von Permutationen Für $n \in \mathbb{N}$ sei S_n die Menge der Permutationen der Menge $\{1, 2, \ldots, n\}$ d.h. S_n ist die Menge der bijektiven Abbildungen $\sigma : \{1, 2, \ldots, n\} \to \{1, 2, \ldots, n\}$. Eine Permutation schreibt man oft auch explizit in der Form

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

Beispielweise ist $\sigma_1 = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix} \in S_3$ die Permutation, die 1 auf 2, 2 auf 3 und 3 auf 1 abbildet während $\sigma_2 = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix} \in S_3$ die Permutation bezeichnet, die 1 und 2 vertauscht und 3 festhält.

Für zwei Permutationen $\sigma, \tau \in S_n$ ist auch die *Hintereinanderausführung* (oder *Komposition*) von τ und σ

$$\sigma \circ \tau : \{1, 2, \ldots, n\} \longrightarrow \{1, 2, \ldots, n\}, k \mapsto \sigma(\tau(k))$$

wieder eine Permutation. Man wendet also auf $k \in \{1, 2, \ldots, n\}$ zunächst τ und auf das Ergebnis $\tau(k)$ dann σ an.

1.3.21 Beispiel: Für $\sigma_1 = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}, \sigma_2 = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix} \in S_3$ gilt

$$\sigma_1 \circ \sigma_2 = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix}, \quad \sigma_2 \circ \sigma_1 = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix}$$

(Für $n \geq 3$ und $\sigma, \tau \in S_n$ ist also im Allgemeinen $\sigma \circ \tau \neq \tau \circ \sigma$.)

1.3.22 Definition: Sei $n \in \mathbb{N}$ und $\sigma \in S_n$.

(a) Ein Paar (i, j) mit $i, j \in \{1, \ldots, n\}$ und $i < j$ heisst Fehlstand von σ, falls $\sigma(i) > \sigma(j)$.

(b) Das Signum (oder Vorzeichen) $\mathrm{sign}(\sigma)$ von σ ist definiert durch

$$\mathrm{sign}(\sigma) = \begin{cases} +1 & \text{falls } \sigma \text{ eine gerade Anzahl von Fehlständen besitzt,} \\ -1 & \text{falls } \sigma \text{ eine ungerade Anzahl von Fehlständen besitzt.} \end{cases}$$

1.3.23 Beispiel: Wir betrachten wieder die beiden Permuationen σ_1, σ_2 aus dem Beispiel 1.3.21. Die Fehlstände der Permutation σ_1 sind die Paare $(1, 3)$ und $(2, 3)$. Also ist $\mathrm{sign}(\sigma_1) = +1$. Dagegen ist $\mathrm{sign}(\sigma_1) = -1$, denn σ_1 hat nur einen Fehlstand (das Paar $(1, 2)$).

Für theoretische Überlegungen ist die folgende Formel zur Berechnung des Signums hilfreich:

1.3.24 Lemma Ist $\sigma \in S_n$ so gilt

$$\mathrm{sign}(\sigma) = \prod_{i<j} \frac{\sigma(j) - \sigma(i)}{j - i}. \tag{5}$$

Dabei durchlaufen i und j in dem Produkt rechts alle Zahlen in $\{1, 2, \ldots, n\}$ mit der Nebenbedingung $i < j$.

Beweis: Für $i, j \in \{1, \ldots, n\}$ mit $i < j$ ist das Vorzeichen des Faktors $\frac{\sigma(j)-\sigma(i)}{j-i}$ negativ oder positiv, je nachdem ob (i, j) ein Fehlstand von σ ist oder nicht. Also ist das Vorzeichen des Produkts $\prod_{i<j} \frac{\sigma(j)-\sigma(i)}{j-i}$ gleich $\mathrm{sign}(\sigma)$. Zum Nachweis der Formel genügt es also zu zeigen dass $\prod_{i<j} \frac{\sigma(j)-\sigma(i)}{j-i}$ Betrag 1 hat, d.h. die Produkte der Beträge der Zähler und der Nenner stimmen überein:

$$\prod_{i<j} |j - i| = \prod_{i<j} |\sigma(j) - \sigma(i)| \tag{6}$$

Das Produkt links kann auch wie folgt beschrieben werden: Für eine 2-elementige Teilmenge M von $\{1, \ldots, n\}$ bezeichne $d(M)$ den Abstand der beiden Elemente, d.h. für $M = \{i, j\}$ mit $i \neq j$ ist $d(M) = |j - i|$. Dann ist $\prod_{i<j} |j - i| = \prod_M d(M)$ wobei M alle 2-elementige Teilmengen von $\{1, \ldots, n\}$ durchläuft. Ebenso gilt für die rechte Seite von (6) die Gleichung $\prod_{i<j} |\sigma(j) - \sigma(i)| = \prod_M d(\sigma(M))$. Dabei ist $\sigma(M)$ das Bild von $M \subseteq \{1, \ldots, n\}$ unter σ (für $M = \{, i, j\}$ bezeichnet $\sigma(M)$ also die Menge $\{\sigma(i), \sigma(j)\}$). Die Gleichung (6) lässt sich also umformulieren zu

$$\prod_M d(M) = \prod_M d(\sigma(M)) \tag{7}$$

wobei M jeweils alle 2-elementigen Teilmengen von $\{1, \ldots, n\}$ durchläuft. Da σ bijektiv ist durchläuft mit M aber auch $\sigma(M)$ alle 2-elementige Teilmengen von $\{1, \ldots, n\}$, d.h. die beiden Produkte in (6) besitzen dieselben Faktoren. $\square$

1.3.25 Lemma Für $\sigma, \tau \in S_n$ gilt

$$\operatorname{sign}(\sigma \circ \tau) = \operatorname{sign}(\sigma) \operatorname{sign}(\tau). \tag{8}$$

Beweis: Für $i, j \in \{1, \ldots, n\}$ mit $i \neq j$ und $\sigma \in S_n$ gilt

$$\frac{\sigma(j) - \sigma(i)}{j - i} = \frac{\sigma(i) - \sigma(j)}{i - j}$$

d.h. der Quotient $\frac{\sigma(j)-\sigma(i)}{j-i}$ hängt nur von der Menge $M\colon = \{i, j\}$ und nicht vom geordneten Paar (i, j) ab. Wir setzen $Q(\sigma, \{i, j\})\colon = \frac{\sigma(j)-\sigma(i)}{j-i}$. Mit (5) gilt dann:

$$\begin{aligned}
\operatorname{sign}(\sigma \circ \tau) &= \prod_{i<j} \frac{\sigma(\tau(j)) - \sigma(\tau(i))}{j - i} \\
&= \prod_{i<j} \frac{\sigma(\tau(j)) - \sigma(\tau(i))}{\tau(j) - \tau(i)} \frac{\tau(i) - \tau(j)}{i - j} \\
&= \left(\prod_{i<j} \frac{\sigma(\tau(j)) - \sigma(\tau(i))}{\tau(j) - \tau(i)} \right) \left(\prod_{i<j} \frac{\tau(i) - \tau(j)}{i - j} \right) \\
&= \left(\prod_M Q(\sigma, \tau(M)) \right) \left(\prod_M Q(\tau, M) \right) \\
&\overset{*}{=} \left(\prod_M Q(\sigma, M) \right) \left(\prod_M Q(\tau, M) \right) = \operatorname{sign}(\sigma) \operatorname{sign}(\tau)
\end{aligned}$$

Dabei durchläuft M in $\prod_M$ wie im Beweis von Lemma 1.3.24 alle 2-elementigen Teilmengen von $\{1, \ldots, n\}$. Die Gleichheit $(*)$ folgt wieder aus der Tatsache, dass mit M auch $\tau(M)$ alle 2-elementigen Teilmengen durchläuft.

Nach diesem Exkurs zu Permutationen kehren wir zu Determinanten zurück. Im Spezialfall $n = 3$ stimmt die folgende Formel mit der Formel von Sarrus überein.

1.3.26 Satz: Sei $A = (a_{ij}) \in M(n \times n, K)$. Dann gilt:

$$\det A = \sum_{\sigma \in S_n} \operatorname{sign}(\sigma) \cdot a_{1\sigma(1)} a_{2\sigma(2)} \cdot \ldots \cdot a_{n\sigma(n)}. \tag{9}$$

Da es genau $n!$ Permutationen der Menge $\{1, \ldots, n\}$ gibt besteht die rechte Seite der Formel (9) aus $n!$ Summanden. Da mit wachsendem n die Fakultät $n!$ rapide steigt, ist die Formel zur Berechnung der Determinante für grosse n ungeeignet.

Zum Beweis von (9) überprüft man, dass der Ausdruck rechts die Eigenschaften (D1)–(D3) aus Satz 1.3.8 besitzt. Wir führen nur den Beweis von (D2) vor, d.h. wir zeigen dass die Summe rechts in (9) für eine $n \times n$-Matrix mit zwei gleichen Zeilen Null ist. Sei also $A = (a_{ij}) \in M(n \times n, K)$ und sei etwa die erste und zweite Zeile von A identisch, d.h. $a_{1j} = a_{2j}$ für $j = 1, \ldots, n$. Mit $\tau \in S_n$ bezeichnen wir die Permutation die 1 und 2 vertauscht und alle anderen Elemente in $\{1, 2, \ldots, n\}$ festhält, d.h.

$$\tau = \begin{pmatrix} 1 & 2 & 3 & \ldots & n \\ 2 & 1 & 3 & \ldots & n \end{pmatrix}.$$

Dann ist $\tau \circ \tau = \mathrm{id}$ die identische Permutation und es gilt $a_{\tau(i)j} = a_{ij}$ für alle $i, j \in \{1, \ldots, n\}$. Da τ nur einen Fehlstand hat (das Paar $(1, 2)$) ist ausserdem $\operatorname{sign}(\tau) = -1$.

Mit σ durchläuft auch $\sigma \circ \tau$ alle Elemente aus S_n (in der Tat gilt $(\sigma \circ \tau) \circ \tau = \sigma$ für alle $\sigma \in S_n$, d.h. die Abbildung $F : S_n \to S_n, F(\sigma) : = \sigma \circ \tau$ ist bijektiv und ihre Umkehrabbildung ist wieder F). Daher können wir im Ausdruck rechts in der Formel (9) die Permutation σ durch $\sigma \circ \tau$ ersetzen und erhalten:

$$\begin{aligned}
\sum_{\sigma \in S_n} \operatorname{sign}(\sigma) \prod_{i=1,\ldots,n} a_{i\sigma(i)} &= \sum_{\sigma \in S_n} \operatorname{sign}(\sigma \circ \tau) \prod_{i=1,\ldots,n} a_{i\sigma(\tau(i))} \\
&= \sum_{\sigma \in S_n} \operatorname{sign}(\sigma \circ \tau) \prod_{i=1,\ldots,n} a_{\tau(i)\sigma(\tau(i))} \\
&\overset{(8)}{=} \sum_{\sigma \in S_n} \operatorname{sign}(\sigma) \operatorname{sign}(\tau) \prod_{i=1,\ldots,n} a_{\tau(i)\sigma(\tau(i))} \\
&= -\sum_{\sigma \in S_n} \operatorname{sign}(\sigma) a_{2\sigma(2)} a_{1\sigma(1)} \cdot a_{3\sigma(3)} \cdot \ldots \cdot a_{n\sigma(n)} \\
&= -\sum_{\sigma \in S_n} \operatorname{sign}(\sigma) \prod_{i=1,\ldots,n} a_{i\sigma(i)}
\end{aligned}$$

d.h. $\sum_{\sigma\in S_n} \operatorname{sign}(\sigma)\prod_{i=1,\dots,n} a_{i\sigma(i)} = 0$. Damit ist gezeigt, dass die Abbildung

$$M(n\times n, K) \to K, A = (a_{ij}) \mapsto \sum_{\sigma\in S_n} \operatorname{sign}(\sigma) \prod_{i=1,\dots,n} a_{i\sigma(i)}$$

die Eigenschaft (D2) besitzt. Der Nachweis der Eigenschaften (D1) und (D3) ist einfacher und sei dem Leser zur Übung überlassen.

2 Gruppen und Körper

2.1 Gruppen

Sei X eine Menge. Eine *Verknüpfung* auf X ist eine Abbildung

$$\star : X\times X \longrightarrow X, (x,y) \mapsto \star(x,y)$$

Die Addition und Multiplikation reeller Zahlen sind z.B. Verknüpfungen auf $\mathbb{R}$. Anstelle von $\star(x,y)$ schreiben wir auch $x\star y$.

2.1.1 Definition: Eine Menge G zusammen mit einer Verknüpfung

$$\star : G\times G \to G, (a,b) \mapsto a\star b$$

heisst Gruppe, wenn die folgenden Axiome erfüllt sind:

(G1) Die Verknüpfung ist assoziativ, d.h. es gilt

$$(a\star b)\star c = a\star(b\star c) \qquad \forall a,b,c\in G.$$

(G2) G besitzt ein neutrales Element, d.h. es gibt ein Element $e\in G$ mit

$$a\star e = e\star a = a \qquad \forall a\in G$$

(G3) Jedes Element von G besitzt ein Inverses, d.h. für alle $a\in G$ existiert ein $b\in G$ mit $a\star b = b\star a = e$.

G heisst kommutativ oder abelsch, wenn zusätzlich gilt:

$$a\star b = b\star a \qquad \forall a\in G.$$

Für die Verknüpfung in einer Gruppe schreibt man häufig auch $a \cdot b$ oder einfach ab (anstelle von $a \star b$). Ist die Verknüpfung als Addition $+$ geschrieben, so setzt man stillschweigend voraus, dass G kommutativ ist.

Aufgrund des Assoziativgesetzes (G1) kann man in mehrfachen Produkten die Klammern weglassen (man schreibt also abc anstelle von $a(bc) = (ab)c$).

Das Inverse von $a \in G$ ist eindeutig bestimmt. Sind nämlich b, b' Elemente in G für die $ab = ba = e$ und $ab' = b'a = e$ gilt, so folgt

$$b = eb = (b'a)b = b'(ab) = b'e = b'$$

Das Inverse von $a \in G$ bezeichnet man mit a^{-1} (bzw. mit $-a$, wenn die Verknüpfung als Addition geschrieben ist).

2.1.2 Bemerkung: Man kann eine Gruppe auch durch folgende (scheinbar schwächere) Axiome charakterisieren:

(G1) $(ab)c = a(bc)$ für alle $a, b, c \in G$.

(G2') Es gibt ein Element $e \in G$ mit $ea = a$ für alle $a, b, c \in G$ (d.h. es gibt ein *linksneutrales* Element).

(G3') Für alle $a \in G$ existiert ein $b \in G$ mit $ba = e$ (d.h. jedes Element besitzt ein *Linksinverses*).

d.h. erfüllt G die Bedingungen (G1), (G2') und (G3') so gelten auch (G1), (G2) und (G3).

Beweis: Sei $e \in G$ ein linksneutrales Element. Dann gilt $ee = e$ (wähle in (G2') speziell $a = e$). Wir folgern jetzt zunächst aus (G1), (G2') und (G3') nachstehende *Kürzungsregel*:

(10) *Für alle $x, y, z \in G$ gilt: $xy = xz \quad \Rightarrow x = y$.*

Multipliziere dazu die Gleichung $xy = xz$ von links mit einem Linksinversen $u \in G$ von x. Wir erhalten:

$$y \overset{(G2')}{=} ey \overset{(G3')}{=} (ux)y \overset{(G1)}{=} u(xy) = u(xz) \overset{(G1)}{=} (ux)z \overset{(G3')}{=} ez \overset{(G2')}{=} z.$$

Wir zeigen jetzt (G2). Für $a \in G$ sei $b \in G$ linksinverses zu a, d.h. $ba = e$. Dann gilt $b(ae) \overset{(G1)}{=} (ba)e = ee = e = ba$ und damit nach (10) (für $x = b, y = ae, z = a$) auch $ae = a$.

Ferner gilt $b(ab) = (ba)b = eb = b \overset{(G2)}{=} be$ und folglich (nach (10)) $ab = e$, d.h. b ist auch rechtsinvers zu a. $\square$

2.1.3 Beispiele: (a) Die Menge der ganzen Zahlen $\mathbb{Z}$ mit der Addition, dem neutralen Element 0 und dem Inversen $-a$ zu a ist eine abelsche Gruppe. Dagegen bildet die Menge der nichtnegativen ganzen Zahlen $\mathbb{N}_0 = \{0, 1, 2, 3, \ldots\}$ mit der Addition keine Gruppe.

(b) $\mathbb{R}$ mit der Addition $+$ und $\mathbb{R} - \{0\}$ mit der Multiplikation sind Gruppen.

(c) Sei $\mathrm{GL}_n(\mathbb{R})$ die Menge aller invertierbaren $n \times n$-Matrizen mit reellen Einträgen. $\mathrm{GL}_n(\mathbb{R})$ bildet zusammen mit der Matrizenmultiplikation eine Gruppe. Das neutrale Element ist E_n und das inverse Element von $A \in \mathrm{GL}_n(\mathbb{R})$ ist die Inverse Matrix A^{-1}. Die Gruppe $\mathrm{GL}_n(\mathbb{R})$ ist nicht abelsch für $n \geq 2$ (vgl. Bemerkung 1.2.7).

(d) Sei $n \in \mathbb{N}$. Die Menge der Permutationen S_n zusammen mit der Hintereinanderausführung $\sigma\tau :\ = \sigma \circ \tau$ als Verknüpfung ist eine Gruppe. Das neutrale Element ist die identische Abbildung $\mathrm{id} : \{1, \ldots, n\} \to \{1, \ldots, n\}, k \mapsto k$. Das Inverse von $\sigma \in S_n$ ist die Umkehrabbildung σ^{-1}. Zum Beispiel ist das Inverse der Permutation $\sigma_1 = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix} \in S_3$ die Permutation $\begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix} \in S_3$ und das Inverse von $\sigma_2 = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix}$ wieder σ_2. Für $n \geq 3$ ist S_n nicht kommutativ (vgl. Beispiel 1.3.21).

2.1.4 Lemma: Sei G eine Gruppe.

(a) Es gelten die folgenden Kürzungsregeln:

$$\begin{aligned} ca = cb &\Rightarrow a = b, \\ ac = bc &\Rightarrow a = b. \end{aligned}$$

(b) Für $a, b \in G$ ist die Gleichung $ax = b$ (bzw. $ya = b$) eindeutig durch ein $x \in G$ (bzw. $y \in G$) lösbar.

Beweis: Zu (a): Die erste Aussage wurde bereits im Beweis von 2.1.2 gezeigt. Die zweite Behauptung zeigt man analog.

Zu (b):

$$\begin{aligned} ax = b &\iff a^{-1}(ax) = a^{-1}b \\ &\iff (a^{-1}a)x = a^{-1}b \\ &\iff ex = a^{-1}b \\ &\iff x = a^{-1}b \end{aligned}$$

$\square$

2.1.5 Definition: Sei G eine Gruppe (mit der Verknüpfung $(a, b) \mapsto ab$). Eine nichtleere Teilmenge H von G heisst Untergruppe, wenn gilt:

(1) $a, b \in H \;\Rightarrow ab \in H$ (d.h. H ist abgeschlossen unter der Verknüpfung).

(2) $a \in H \;\Rightarrow a^{-1} \in H$ (d.h. mit jedem Element in H liegt auch das Inverse in H).

In jeder Untergruppe $H \subseteq G$ liegt das neutrale Element. Wähle dazu ein beliebiges Element $a \in H$ (nach Voraussetzung ist $H \neq \emptyset$). Nach (2) ist dann $a^{-1} \in H$ und mit (1) folgt $e = aa^{-1} \in H$.

2.1.6 Bemerkung: (a) Eine nichtleere Teilmenge H von G ist genau dann eine Untergruppe, wenn gilt:

$$a, b \in H \Longrightarrow ab^{-1} \in H$$

(b) Ist H Untergruppe von G so ist H mit der Verknüpfung aus G wieder eine Gruppe.

2.1.7 Beispiele: (a) Jede Gruppe besitzt die beiden trivialen Untergruppen $\{e\}$ und G.

(b) Für eine natürliche Zahl n ist die Menge aller Vielfachen $n\mathbb{Z} := \{mn \mid m \in \mathbb{Z}\}$ eine Untergruppe von $\mathbb{Z}$ (mit der Addition als Verknüpfung).

(c) Die Menge der invertierbaren oberen Dreiecksmatrizen mit reellen Einträgen ist eine Untergruppe von $\mathrm{GL}_n(\mathbb{R})$.

2.2 Körper

Ein Körper ist eine Menge mit zwei Verknüpfungen (die man in Anlehnung an die reellen Zahlen als Addition und Multiplikation bezeichnet) die gewisse Axiome erfüllen und durch das *Distributivgesetzt* miteinander in Beziehung stehen. Die genaue Definition lautet:

2.2.1 Definition: Ein Körper K ist eine Menge zusammen mit zwei Verknüpfungen

$$+ : K \times K \longrightarrow K, (a, b) \mapsto a + b \qquad \text{(Addition)}$$

und

$$\cdot : K \times K \to K, (a,b) \mapsto ab \qquad \text{(Multiplikation)}$$

so dass die folgenden Axiome erfüllt sind:

(1) $a + (b + c) = (a + b) + c$ für alle $a, b, c \in K$.

(2) Es existiert ein Element $0 = 0_K \in K$, so dass

$$0 + a = a + 0 = a \qquad \forall\, a \in K.$$

(3) Für jedes $a \in K$ gibt es ein $-a \in K$, so dass

$$a + (-a) = 0.$$

(4) $a + b = b + a$ für alle $a, b \in K$.

(5) $a(bc) = (ab)c$ für alle $a, b, c \in K$.

(6) Es existiert ein Element $1 = 1_K \in K$, $1 \neq 0$, so dass

$$1a = a1 = a \qquad \forall\, a \in K.$$

(7) Für jedes $a \in K$, $a \neq 0$ gibt es ein $a^{-1} \in K$ mit

$$aa^{-1} = 1.$$

(8) $ab = ba$ für alle $a, b \in K$.

(9) (Distributivgesetzt) $a(b + c) = ab + ac$ für alle $a, b, c \in K$.

2.2.2 Bemerkung: Die Axiome (1)–(9) können wie folgt zusammengefasst werden:

(i) K mit der Addition $+$ ist eine abelsche Gruppe.

(ii) $K - \{0\}$ mit der Multiplikation ist eine abelsche Gruppe.

(iii) Es gilt das Distributivgesetzt: $a(b + c) = ab + ac, \quad \forall a, b, c \in K$.

2.2.3 Beispiele: Die reellen Zahlen $\mathbb{R}$ und die rationalen Zahlen $\mathbb{Q}$ bilden jeweils einen Körper. Dagegen sind die ganzen Zahlen $\mathbb{Z}$ kein Körper, denn $\mathbb{Z}$ erfüllt zwar die Axiome (1)–(6), (8) und (9) aber nicht (7).

Ein weiteres wichtiges Beispiel ist der Körper der komplezen Zahlen der im nächsten Abschnitt eingeführt wird. Zunächst soll aber noch gezeigt werden, dass es für jede Primzahl p einen Körper mit p Elementen gibt.

Teilbarkeit und Kongruenzrelation Seien a, b ganze Zahlen mit $a \neq 0$. Dann heisst b ein *Vielfaches* von a wenn es eine ganze Zahl q gibt mit $b = qa$. Man sagt dann auch a teilt b (oder a ist ein Teiler von b) und benutzt dafür die Notation $a \mid b$. Falls a nicht b teilt so schreibt man $a \nmid b$.

2.2.4 Lemma: (Division mit Rest) Seien $m, n \in \mathbb{Z}$ gegeben und es gelte $n \neq 0$. Dann gibt es eindeutig bestimmte ganze Zahlen q und r mit der Eigenschaft

$$m = qn + r, \qquad r \in \{0, 1, \dots, |n| - 1\}. \tag{11}$$

Die Zahl r heisst der *Rest von m bei der Division durch n.*

Beweis: Sei zunächst $n \geq 0$ und sei $q \colon = [\frac{m}{n}]$ die grösste ganze Zahl $\leq \frac{m}{n}$, d.h. $q \in \mathbb{Z}$ und es gelten die Ungleichungen $q \leq \frac{m}{n} < q + 1$. Für $r \colon = m - qn \in \mathbb{Z}$ gilt dann also

$$r = n(\frac{m}{n} - q) \geq 0 \quad \text{und} \quad r = n(\frac{m}{n} - q) < n(q + 1 - q) = n$$

d.h. $r \in \{0, 1, \dots, n-1\}$. Damit ist die Existenz der Darstellung (11) im Fall $n > 0$ gezeigt.

Sei jetzt $n < 0$. Dann können wir das obige Argument auf $|n| = -n$ anwenden, d.h. wir teilen m zunächst durch $|n|$ mit Rest

$$m = \widetilde{q}|n| + \widetilde{r}, \quad \widetilde{r} \in \{0, 1, \dots, -n - 1\}$$

Setzen wir jetzt $q \colon = -\widetilde{q}$ und $r \colon = \widetilde{r}$ so erhalten wir $m = qn + r$ mit $0 \leq r \leq -n - 1 = |n| - 1$.

Es bleibt zu zeigen, dass die beiden Zahlen q und r in (11) eindeutig bestimmt sind. Sei (q', r') ein weiteres Paar von ganzen Zahlen, dass (11) erfüllt. Aus $qn + r = m = q'n + r'$ folgt $n(q - q') = r' - r$, d.h. $r' - r$ ist ein Vielfaches von $|n|$. Wegen $r, r' \in \{0, 1, \dots, |n| - 1\}$ gilt andererseits

$$-|n| < -r = 0 - r \leq r' - r < |n| - r \leq |n|$$

d.h $|r' - r| < |n|$. Damit folgt $r' = r$ und wegen $qm + r = q'n + r'$ und $n \neq 0$ auch $q = q'$. $\square$

2.2.5 Definition: Sei $n \in \mathbb{Z}$ mit $n \neq 0$ und $a, b \in \mathbb{Z}$. Man sagt, dass a kongruent zu b ist modulo m, wenn $a - b$ durch m teilbar ist. Dafuer benutzt man die Notation $a \equiv b \mod m$.

Zum Beispiel ist $137 \equiv 11 \mod 7$, denn $126 = 7 \cdot 18$.

2.2.6 Satz: (Rechenregeln für die Kongruenzrelation). Sei $n \in \mathbb{Z}$ mit $n \neq 0$. Für alle $a, b, c, d \in \mathbb{Z}$ gilt:

(1) $a \equiv a$ für alle $a \in \mathbb{Z}$.

(2) Für alle $a, b \in \mathbb{Z}$ gilt: $a \equiv b \Longrightarrow b \equiv a$.

(3) Für alle $a, b, c \in \mathbb{Z}$ gilt:

$$a \equiv b \text{ und } a \equiv c \Longrightarrow a \equiv c.$$

(4) Für alle $a, b, c \in \mathbb{Z}$ gilt:

$$a \equiv b \text{ und } c \equiv d \Longrightarrow a + b \equiv c + d.$$

(5) Für alle $a, b, c \in \mathbb{Z}$ gilt:

$$a \equiv b \text{ und } c \equiv d \Longrightarrow ab \equiv cd.$$

Dabei steht $x \equiv y$ immer abkürzend für $x \equiv y \mod n$.

Beweis: Wir überprüfen nur (5) und überlassen den Nachweis der anderen Behauptungen dem Leser. Wegen $a \equiv b$ und $c \equiv d$ gibt es ganze Zahlen q_1, q_2 mit $a - b = q_1 n$ und $c - d = q_2 n$. Dann gilt mit $q_3 := aq_2 + dq_1$:

$$ac - bd = a(c - d) + d(a - b) = aq_2 n + dq_1 n = q_3 n$$

d.h. $ab \equiv cd \mod n$ wie behauptet. $\square$

Endliche Körper Für $n \in \mathbb{N}$ sei $\mathbb{Z}_n$ die Menge $\{0, 1, \ldots, n-1\}$ mit der Addition:

$$i \oplus j = \text{Rest von } i + j \text{ bei der Division durch } n$$

und der Multiplikation

$$i \odot j = \text{Rest von } i \cdot j \text{ bei der Division durch } n,$$

d.h. für $i, j \in \{0, 1, \ldots, n-1\}$ sind $i \oplus j \in \{0, 1, \ldots, n-1\}$ und $i \cdot j \in \{0, 1, \ldots, n-1\}$ durch die Kongruenzen

$$i \oplus j \equiv i + j \mod n, \qquad i \odot j \equiv i \cdot j \mod n$$

charkterisiert. Für $n = 5$ erhält man z.B. die folgende *Verknüpfungstafeln* für $\oplus$ und $\odot$:

$\oplus$	0	1	2	3	4
0	0	1	2	3	4
1	1	2	3	4	0
2	2	3	4	0	1
3	3	4	0	1	2
4	4	0	1	2	3

$\odot$	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

Man überprüft leicht anhand der Tafeln, dass $\mathbb{Z}_5$ ein Körper ist. Z.B. existieren die Inversen der Zahlen $1, 2, 3$ und 4 bzgl. der Multiplikation (das sind nämlich $1, 3, 2$ und 4). Allgemein gilt:

2.2.7 Satz: Die Menge $\mathbb{Z}_n = \{0, 1, \ldots, n-1\}$ mit der Addition $\oplus$ und der Multiplikation $\odot$ ist genau dann ein Körper, wenn n eine Primzahl ist.

Dabei heisst eine natürliche Zahl $p > 1$ Primzahl, wenn für alle $a, b \in \mathbb{Z}$ gilt:

$$p \mid ab \quad \Rightarrow \quad p \mid a \quad \text{oder} \quad p \mid b. \tag{12}$$

Man kann leicht zeigen, dass $p \in \mathbb{N}, p > 1$ genau dann eine Primzahl ist, wenn es nur die Teiler $\pm p, \pm 1$ besitzt.

Beweis von Satz 2.2.7: Mit Hilfe von Satz 2.2.6 kann man sich leicht überlegen, dass $\mathbb{Z}_n$ immer die Bedingungen (1)–(6), (8) und (9) in Definition 2.2.2 erfüllt. Wir müssen also nur zeigen, dass jedes $i \in \{0, 1, \ldots, n-1\}$ genau dann ein Inverses (bzgl. $\odot$) besitzt, wenn n prim ist.

Für $i \in \mathbb{Z}_n$ bezeichne $\ominus i$ das Inverse von i bzgl. $\oplus$. Für $i, j \in \mathbb{Z}_n$ schreiben wir abkürzend $i \ominus j$ für den Ausdruck $i \oplus (\ominus j)$. Es gilt also

$$i \ominus j = \text{Rest von } i - j \text{ bei der Division durch } n.$$

Wir nehmen zunächst an, dass n eine Primzahl ist. Dann gilt für $i, j \in \mathbb{Z}_n$:

$$(13) \qquad i \odot j = 0 \quad \Rightarrow \quad i = 0 \quad \text{oder} \quad j = 0$$

Aus $i \odot j = 0$ folgt nämlich $i \cdot j \equiv i \odot j = 0 \mod n$, d.h. $n \mid i \cdot j$ und damit wegen (12) $i = 0$ oder $j = 0$.

Wir wollen jetzt zeigen dass ein fest gewähltes $i \in \mathbb{Z}_n - \{0\}$ ein Inverses bzgl. der Multiplikation besitzt. Betrachte dazu die $(n-1)$-Produkte $i \odot 1$, $i \odot 2, \ldots, i \odot (n-1)$. Nach (13) sind sie alle von 0 verschieden. Sie sind aber auch untereinander verschieden, denn aus $i \odot j_1 = i \odot j_2$ folgt (mit Hilfe des Distributivgesetzt):

$$i \odot (j_1 \ominus j_2) = (i \odot j_1) \ominus (i \odot j_2) = 0$$

und damit wegen (13) auch $j_1 \ominus j_2 = 0$ also $j_1 = j_2$.

Wir haben gezeigt, dass die $n-1$ Zahlen $i \odot 1, i \odot 2, \ldots, i \odot (n-1)$ in $\mathbb{Z}_n$ alle paarweise verschieden sind und $\neq 0$. Also muss unter ihnen die 1 vorkommen, d.h. es gibt ein $j \in \mathbb{Z}_n - \{0\}$ mit $i \odot j = 1$. Also ist $\mathbb{Z}_n$ ein Körper ist.

Wir müssen jetzt noch die Umkehrung zeigen, d.h. falls $\mathbb{Z}_n$ ein Körper ist, so ist n eine Primzahl. Seien $a, b \in \mathbb{Z}$ mit $p \mid ab$. Sei i der Rest von $a \in \mathbb{Z}_n$ und $j \in \mathbb{Z}_n$ der Rest von b bei der Division durch n. Nach 2.2.6 (5) gilt

$$i \odot j \equiv i \cdot j \equiv a \cdot b \equiv 0 \mod n$$

und damit (wegen $0 \leq i \odot j \leq n-1$) $i \odot j = 0$. Angenommen $n \nmid a$ also $i \neq 0$. Dann besitzt i ein Inverses $k \in \mathbb{Z}_n$ bzgl. $\odot$, d.h. es gilt $k \odot i = 0$.

$$\Rightarrow j = 1 \odot j = k \odot i \odot j = k \odot 0 = 0 \quad \Rightarrow \quad b \equiv j \equiv 0 \mod n \quad \Rightarrow \quad n \mid b$$

Das beweist, dass n prim ist. $\square$

Komplexe Zahlen Die komplexen Zahlen sind aus dem Wunsch heraus entstanden den Zahlbereich der reellen Zahlen so zu vergrössern, dass die Gleichung $x^2 + 1 = 0$ ein Lösung besitzt. Eine Lösung dieser Gleichung wird mit i bezeichnet. Komplexen Zahlen sind formale Ausdrücke der Form $x + iy$ mit denen man “genauso rechnet” wie mit reellen Zahlen, wobei man berücksichtigt, dass $i^2 = -1$. Die formale Definition der komplexen Zahlen ist wie folgt:

2.2.8 Definition: Die komplexen Zahlen $\mathbb{C}$ sind die Menge $\mathbb{R}^2 = \mathbb{R} \times \mathbb{R}$, zusammen mit der Addition

$$(x, y) + (u, v) := (x + u, y + v)$$

und der Multiplikation

$$(x, y) \cdot (u, v) := (xu - yv, xv + yu).$$

2.2.9 Satz: $\mathbb{C}$ ist ein Körper.

Beweis: Man rechnet leicht nach, dass die Addition (A) (bzw. die Multiplikation (M)) assoziativ und kommutativ sind. Das Distributivgesetz gilt wegen:

$$\begin{aligned}
&(x, y) \cdot ((u_1, v_1) + (u_2, v_2)) = \\
=&(x, y) \cdot (u_1 + u_2, v_1 + v_2) = \\
=&(x(u_1 + u_2) - y(v_1 + v_2), x(v_1 + v_2) + y(u_1 + u_2)) = \\
=&(xu_1 - yv_1 + xu_2 - yv_2, xv_1 + yu_1 + xv_2 + yu_2) = \\
=&(xu_1 - yv_1, xv_1 + yu_1) + (xu_2 - yv_2, xv_2 + yu_2) = \\
=&(x, y) \cdot (u_1, v_1) + (x, y) \cdot (u_2, v_2).
\end{aligned}$$

Das neutrale Element bzgl. der Addition ist $(0, 0)$ und das Inverse von (x, y) ist $(-x, -y)$. Das neutrale Element bzgl. der Multiplikation ist $(1, 0)$, denn

$$(x, y) \cdot (1, 0) = (x\,1 - y\,0, x\,0 + y\,1) = (x, y) = (1, 0) \cdot (x, y)$$

für alle komplexe Zahlen (x, y). Schliesslich bleibt noch zu zeigen, dass eine komplexe Zahl $(x, y) \neq (0, 0)$ ein Inverses bzgl. der Multiplikation besitzt. Wegen $x \neq 0$ oder $y \neq 0$ ist $x^2 + y^2 > 0$. Da

$$(x, y) \cdot \left(\frac{x}{x^2 + y^2}, \frac{-y}{x^2 + y^2} \right) = \left(\frac{x^2}{x^2 + y^2} + \frac{y^2}{x^2 + y^2}, \frac{-xy}{x^2 + y^2} + \frac{yx}{x^2 + y^2} \right) = (1, 0)$$

ist $(x, y)^{-1} :\, = \left(\frac{x}{x^2+y^2}, \frac{-y}{x^2+y^2} \right)$ das Inverse von (x, y) bzgl. der Multiplikation.

Wir haben also gesehen, dass $\mathbb{C}$ den Körperaxiomen 2.2.1 (1) – (9) genügt. $\square$

2.2.10 $\mathbb{R}$ als Unterkörper von $\mathbb{C}$: Für komplexe Zahlen der Gestalt $(x, 0)$ gilt:

$$\begin{aligned}
(x, 0) + (y, 0) &= (x + y, 0) \\
(x, 0) \cdot (y, 0) &= (xy, 0)
\end{aligned}$$

Die komplexen Zahlen der Gestalt $(x,0),(y,0)$ werden also wie die entsprechenden reellen Zahlen x,y addiert und multipliziert. Wir wollen im Folgenden nicht mehr zwischen $x \in \mathbb{R}$ und der komplexen Zahl $(x,0)$ unterscheiden. Wir schreiben also einfach x für $(x,0)$ und fassen auf diese Weise $\mathbb{R}$ als Teilmenge von $\mathbb{C}$ auf. Insbesondere schreiben wir einfach 0 für $(0,0)$ und 1 statt $(1,0)$. Für jede komplexe Zahl gilt dann $z+0=z$, $z \cdot 1 = z$.

Da für $\mathbb{R} \subset \mathbb{C}$ die Addition und Multiplikation in $\mathbb{C}$ mit der gewöhnlichen Addition und Multiplikation übereinstimmt, spricht man davon, dass $\mathbb{R}$ ein *Unterkörper* von $\mathbb{C}$ ist.

2.2.11 Bemerkung: Wenn man sehr genau sein möchte, kann man den Zusammenhang zwischen $\mathbb{R}$ und $\mathbb{C}$ wie folgt beschreiben: Die Abbildung

$$\begin{aligned} \iota : \mathbb{R} &\to \widetilde{\mathbb{R}} := \{(x,0) \mid x \in \mathbb{R}\} \\ x &\mapsto (x,0) \end{aligned}$$

ist ein *Isomorphismus*, d.h. ι ist bijektiv und es gilt: $\iota(x+y) = \iota(x) + \iota(y)$, $\iota(xy) = \iota(x)\iota(y)$ für alle $x,y \in \mathbb{R}$. Also ist $\mathbb{R}$ zum Unterkörper $\widetilde{\mathbb{R}}$ von $\mathbb{C}$ *isomorph* (d.h. $\mathbb{R}$ und $\widetilde{\mathbb{R}}$ sind gleichstrukturiert).

2.2.12 Die imaginäre Einheit: Darunter versteht man die komplexe Zahl $i := (0,1)$.

Es gilt: $i^2 = i \cdot i = (0,1)(0,1) = (-1,0) = -1$, d.h. i und $-i$ sind die Lösungen der Gleichung $z^2 + 1 = 0$.

Für eine komplexe Zahl $z = (x,y)$ gilt:

$$z = (x,0) + (0,y) = (x,0) + (0,1)(y,0) = x + iy$$

Dies ist die *kartesische Darstellung* von z. Die reellen Zahlen x und y heissen der Real- bzw. der Imaginärteil von z (in Zeichen: $x := \mathrm{Re}(z)$, $y := \mathrm{Im}(z)$). Man kann eine komplexe Zahl z also geometrisch als einen Punkt der Ebene $\mathbb{R}^2$ mit x-Koordinate $\mathrm{Re}(z)$ und y-Koordinate $\mathrm{Im}(z)$ deuten. Man spricht deshalb auch von der *komplexen Zahlenebene*.

2.2.13 Die komplexe Konjugation: Für eine komplexe Zahl $z = (x,y) = x + iy$ setzt man

$$\overline{z} = (x,-y) = x - iy$$

Es gilt: $z\overline{z} = (x+iy)(x-iy) = x^2 - i^2y^2 = x^2 + y^2$ d.h. $z\overline{z}$ ist reell und > 0 falls $z \neq 0$.

2.2.14 Definition: Für $z \in \mathbb{C}$ heisst $|z| := \sqrt{z\overline{z}}$ der Betrag von z.

Geometrisch $|z|$ ist der Abstand des Punktes $z = x + iy = (x, y)$ vom Ursprung (oder anders ausgedrückt: $|z|$ ist die Länge der Verbindungsstrecke von $(0, 0)$ zu (x, y)). Das ergibt sich sofort aus dem Satz des Pythagoras.

Es gelten die folgenden Rechenregeln:

(1) $\mathrm{Re}\,(z) = \frac{1}{2}(z + \overline{z})$, $\mathrm{Im}\,(z) = \frac{1}{2i}(z - \overline{z})$;

(2) $z = \overline{z} \iff z \in \mathbb{R}$;

(3) $\overline{z + w} = \overline{z} + \overline{w}$, $\overline{zw} = \overline{z}\overline{w}$;

(4) $\overline{\overline{z}} = z$;

(5) $|\overline{z}| = |z|$;

(6) $|z \cdot w| = |z|\,|w|$;

(7) $|(x, 0)| = |x| \quad \forall\, x \in \mathbb{R}$;

(8) $|\mathrm{Re}(z)| \leq |z|$, $|\mathrm{Im}(z)| \leq |z|$;

(9) $|z + w| \leq |z| + |w|$ (Dreiecksungleichung).

Wir beweisen nur (9) (der Nachweis von (1)–(8) sei als Übungsaufgabe empfohlen). Geometrisch ergibt sich (9) sofort aus der Tatsache, dass in einem Dreieck die Summe der Längen von zweier Seiten immer grösser als die Länge der Dritten ist (das wendet man auf das Dreieck in der komplexen Zahlenebene mit den Eckpunkten $0, z, z + w$ an). Ein algebraischer Beweis verläuft wie folgt:

$$|z + w|^2 = (z + w)(\overline{z + w}) \overset{(3)}{=} (z + w)(\bar{z} + \bar{w}) =$$

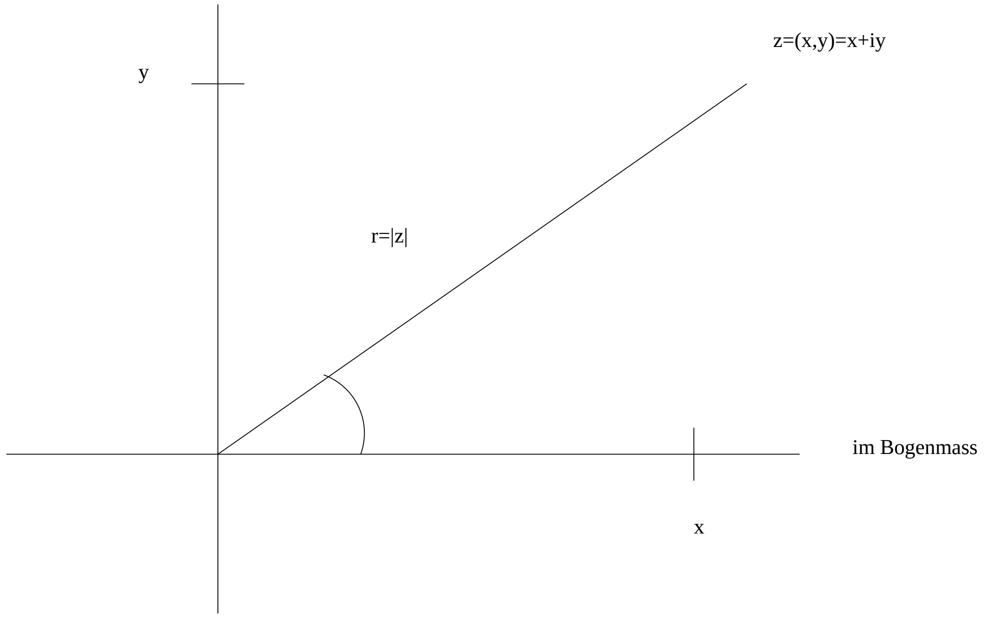
$$= z\bar{z} + z\bar{w} + w\bar{z} + w\bar{w} \overset{(3)}{=} |z|^2 + z\bar{w} + \overline{z\bar{w}} + |w|^2 =$$

$$\overset{(1)}{=} |z|^2 + 2\mathrm{Re}(z\bar{w}) + |w|^2 \overset{(8)}{\leq} |z|^2 + 2|z\bar{w}| + |w|^2 =$$

$$\overset{(5),(6)}{=} |z|^2 + 2|z|\,|w| + |w|^2 = (|z| + |w|)^2$$

Es folgt $|z + w| \leq |z| + |w|$. □

Darstellung der komplexen Zahlen in Polarkoordinaten. Jede komplexe Zahl $z = x + iy \neq 0$ lässt sich in der Form $z = r(\cos\varphi + i\sin\varphi)$ mit eindeutig bestimmten $r \in \mathbb{R}, r > 0$, $\varphi \in [0, 2\pi)$. Dies ist die Darstellung von z in *Polarkoordinaten*. Dabei ist $r = |z|$ der Betrag von z. Dabei ist φ der Winkel zwischen der positiven Richtung der x-Achse und dem Vektor (x, y). Der Winkel φ wird als das Argument von z bezeichnet (in Zeichen $\varphi = \mathrm{Arg}\,(z)$).



Die folgenden Formeln geben die Umrechnung zwischen den kartesischen und Polarkoordinaten von z an:

$$\begin{array}{lll} x = r\cos(\varphi) & r = \sqrt{x^2+y^2} & \\ y = r\sin(\varphi) & \tan\varphi = \frac{y}{x} & \text{(Quadranten beachten).} \end{array}$$

2.2.15 Beispiele: (a) Die Darstellung von $z = \frac{1}{2} + \frac{\sqrt{3}}{2}i$ in Polarkoordinaten ist $\cos\left(\frac{\pi}{3}\right) + i\sin\left(\frac{\pi}{3}\right)$, d.h. $r = 1$ und $\mathrm{Arg}\,(z) = \frac{\pi}{3}$.

(b) Sind umgekehrt die Polarkoordinaten von z, gegeben durch $r = 2$ und $\varphi = \frac{2\pi}{3}$, so ergibt sich für die kartesische Darstellung $z = x + iy$:

$$x = 2\cos\left(\frac{2\pi}{3}\right) = -1, \quad y = 2\sin\left(\frac{2\pi}{3}\right) = \sqrt{3}.$$

In der der Polorkoordinatendarstellung lässt sich das Produkt zweier komplexer Zahlen einfacher beschreiben als in der kartesischen Darstellung. Sind

$z_1 = r_1(\cos\varphi_1 + i\sin\varphi_1)$, $z_2 = r_2(\cos\varphi_2 + i\sin\varphi_2)$ zwei komplexe Zahlen $\neq 0$ so gilt:

$$z_1 z_2 = r_1 r_2(\cos(\varphi_1 + \varphi_2) + i\sin(\varphi_1 + \varphi_2)$$

Das folgt sofort aus den, aus der Schulmathematik bekannten, *Additionstheoremen*:

$$\cos(\varphi_1 + \varphi_2) = \cos\varphi_1\cos\varphi_2 - \sin\varphi_1\sin\varphi_2,$$
$$\sin(\varphi_1 + \varphi_2) = \sin\varphi_1\cos\varphi_2 + \cos\varphi_1\sin\varphi_2.$$

Also gilt: Das Produkt zweier komplexer Zahlen $z_1 z_2$ erhält man, indem man ihre Beträge multipliziert und ihre Argumente addiert (bis auf ein Vielfaches von 2π, d.h. $\operatorname{Arg}(z_1 z_2) = \operatorname{Arg}(z_1) + \operatorname{Arg}(z_2) + 2\pi k$ für eine ganze Zahl k).

2.2.16 Beispiele: (a) Sei $z = \frac{1}{2} + \frac{\sqrt{3}}{2}i = \cos\left(\frac{\pi}{3}\right) + i\sin\left(\frac{\pi}{3}\right)$. Dann ist

$$z^6 = \cos\left(6\frac{\pi}{3}\right) + i\sin\left(6\frac{\pi}{3}\right) = 1.$$

Entsprechend gilt $(\frac{\sqrt{2}}{2}+\frac{\sqrt{2}}{2}i)^8 = 1$, da $|\frac{\sqrt{2}}{2}+\frac{\sqrt{2}}{2}i| = 1$ und $\operatorname{Arg}((\frac{\sqrt{2}}{2}+\frac{\sqrt{2}}{2}i) = \frac{\pi}{4}$.

In der Polkoordinatendarstellung ist das Wurzeln ziehen relativ einfach. Es gilt:

2.2.17 Satz: Sei $z \in \mathbb{C}$, $z \neq 0$ eine komplexe Zahl. Dann hat die Gleichung

$$w^n = z \tag{14}$$

genau n verschiedene Lösungen. Ist $z = r(\cos\varphi + i\sin\varphi)$ die Darstellung von z in Polarkoordinaten, so sind

$$w_k = \sqrt[n]{r}(\cos\frac{\varphi + 2\pi k}{n} + i\sin\frac{\varphi + 2\pi k}{n}), \quad k = 0, 1, \ldots, n-1$$

die verschiedenen Lösungen.

Dabei ist $\sqrt[n]{r}$ die eindeutig bestimmte n-te positive reelle Wurzel von reellen Zahl $r > 0$.

Beweis: Ist $w = t(\cos\psi + i\sin\psi)$ eine Lösung der Gleichung (14), so gilt

$$t^n = r, \qquad n\psi = \varphi + 2\pi k$$

für ein gewisses $k \in \mathbb{Z}$. Wegen $\psi, \varphi \in [0, 2\pi)$ ist

$$2\pi k = n\psi - \varphi \leq n\psi < 2\pi n,$$

und

$$2\pi k = n\psi - \varphi \geq -\varphi > -2\pi$$

also $k \in \{0, 1, \ldots, n-1\}$.

Umgekehrt sind die $w_k := \sqrt[n]{r}\,(\cos\frac{\varphi+2\pi k}{n} + i\,\sin\frac{\varphi+2\pi k}{n})$ für $k = 0, 1, \ldots, n-1$ Lösungen von (14), da $|w_k^n| = \sqrt[n]{r}^n = r$ und

$$\operatorname{Arg}(w_k^n) \doteq n\operatorname{Arg}(w_k) \doteq \varphi + 2k\pi \doteq \varphi$$

wobei $\doteq$ bedeutet, dass eine Gleichheit bis auf ein ganzzahliges Vielfaches von 2π vorliegt. Damit folgt $\operatorname{Arg}(w_k^n) = \operatorname{Arg}(z)$, also $w_k^n = z$. □

2.2.18 Bemerkung: Üblicherweise versteht man unter der Polarkoordinatendarstellung einer komplexen Zahl $z \in \mathbb{C}$ $z \neq 0$, eine Darstellung der Form

$$z = r\,e^{i\varphi}, \qquad \text{mit } r \in \mathbb{R}, r > 0, \varphi \in [0, 2\pi).$$

Dabei ist $e = 2,718\ldots$ die *Eulersche Zahl*.

Diese ist äquivalent zu der Darstellung $z = r(\cos\varphi + i\sin\varphi)$, denn es gilt die *Eulersche Formel*:

$$e^{i\varphi} = \cos\varphi + i\,\sin\varphi.$$

Dazu ist aber zunächst einmal die Bedeutung des Ausdrucks $e^{i\varphi}$ zu klären. Aus der Schulmathematik ist die folgende Formel für die Exponentialfunktion $\exp(x) := e^x$ bekannt:

$$e^x = \exp(x) = \sum_{n=0}^{\infty} \frac{x^n}{n!}$$

Für eine komplexe Zahl z definiert man daher

$$e^z = \exp(z) := \sum_{n=0}^{\infty} \frac{z^n}{n!}$$

Dass man diese Reihe überhaupt bilden (d.h. dass sie konvergiert) folgt aus dem Quotientenkriterium, dass auch für unendliche Reihen aus komplexen Zahlen gilt.

Nullstellen von Polynomen. Sei K ein unendlicher Körper (etwa $K = \mathbb{R}$ oder $\mathbb{C}$).

2.2.19 Definition: Ein Polynom (über K) ist eine Funktion der Form

$$f : K \to K, x \mapsto f(x) = a_n x^n + \ldots + a_1 x + a_0 = \sum_{k=0}^{n} a_k x^k$$

mit $a_0, a_1, \ldots, a_n \in K$. Ist $a_n \neq 0$, so heisst n der Grad des Polynoms (in Zeichen: $n = \deg f$). Sind alle $a_k = 0$, so heißt f das Nullpolynom, dem wir den Grad $-\infty$ zuordnen.

Mit $K[x]$ bezeichnen wir die Menge aller Polynome über K.

2.2.20 Bemerkung: Wir werden später zeigen (Folgerung 2.2.25 unten), dass die $a_0, a_1, \ldots, a_n \in K$ durch f eindeutig bestimmt sind (dabei ist entscheidend, dass K als unendlich vorausgesetzt wurde). Sie heissen die Koeffizienten von f.

Summen und Produkte von Polynomen sind wieder Polynome. Das Produkt der Polynome

$$\begin{aligned} f(x) &= a_n x^n + \ldots + a_1 x + a_0 \\ g(x) &= b_m x^m + \ldots + b_1 x + b_0 \end{aligned}$$

ist das Polynom

$$f(x)g(x) = c_{m+n} x^{m+n} + \ldots + c_1 x + c_0$$

mit

$$c_k = \sum_{r+s=k} a_r b_s \qquad k = 0, 1, \ldots, m+n.$$

2.2.21 Lemma: (Polynomdivision): Sei $g(x)$ ein Polynom $\neq 0$. Dann gibt es zu jedem Polynom $f(x)$ eindeutig bestimmte Polynome $q(x)$ und $r(x)$ mit

$$f(x) = g(x) \cdot q(x) + r(x)$$

und $\deg r(x) < \deg g(x)$.

Beweis der Existenz: Ist $\deg f < \deg g$ so ist $f = g \cdot 0 + f$ eine Zerlegung der gesuchten Art.

Im anderen Fall sei $n := \deg f \geq m := \deg g$, $f(x) = a_n x^n + \ldots + a_1 x + a_0$, $g(x) = b_m x^m + \ldots + b_1 x + b_0$. Dann ist

$$f(x) - a_n b_m^{-1} x^{n-m} g(x) =: f_1(x)$$

ein Polynom vom Grad $n_1 < n$. Ist $n_1 \geq m$ so subtrahieren wir von f_1 wieder ein Vielfaches von g, so dass die Differenz ein Polynom vom Grad $n_2 < n_1$ ist. So fortfahrend erhält man schliesslich ein Polynom r, dessen Grad $< m$ ist. Mit geeignetem q gilt also $f - gq = r$.

Eindeutigkeit: Ist $f = gq' + r'$ eine weitere derartige Zerlegung mit $q \neq q'$, so folgt:

$$g(q - q') = r - r' \Rightarrow \deg\, g(q - q') = \deg\,(r - r') < m,$$

ein Widerspruch. □

2.2.22 Beispiel: Sei $f(x) = 3x^3 + 2x + 1$ und $g(x) = x^2 - x$. Dann gilt

$$f(x) = g(x)(3x + 3) + (5x + 1)$$

d.h. $f = gq + r$ mit $q(x) = 3x + 3$ und $r(x) = 5x + 1$.

2.2.23 Lemma: Sei $f(x)$ ein Polynom über K. Ein Element $\alpha \in K$ ist genau dann Nullstelle von $f(x)$, wenn $f(x)$ durch $x - \alpha$ teilbar ist, d.h. es gibt ein Polynom $q(x)$ vom Grad $\deg\,(g) = \deg\,(f) - 1$ mit

$$f(x) = (x - \alpha)q(x)$$

Beweis: Nach dem Satz über die Division mit Rest gibt es $q(x), r(x) \in K[x]$ mit

$$f(x) = (x - \alpha)q(x) + r(x) \text{ und } \deg\, r(x) < \deg\,(x - \alpha) = 1$$

also $r(x) = c$, $c \in K$. Folglich gilt:

$$f(\alpha) = 0 \iff r(\alpha) = q(\alpha)(\alpha - \alpha) + r(\alpha) = 0 \iff c = 0.$$

□

2.2.24 Folgerung: Ein Polynom $\neq 0$ vom Grad n hat höchstens n verschiedene Nullstellen.

Beweis durch Induktion über n: Der Fall $n = 1$ ist klar.
$n \to n+1$: Sei f ein Polynom vom Grad $n+1$. Falls f keine Nullstelle hat ist nichts zu zeigen. Andernfalls sei α eine Nullstellen von f. Nach dem obigen Lemma gibt es ein Polynom g vom Grad n mit

$$f(x) = (x - \alpha_1)g(x).$$

Nach der Induktionsvoraussetzung hat g höchstens n Nullstellen. Damit hat f höchstens $n+1$ Nullstellen, denn jede solche ist entweder $= \alpha$ oder Nullstelle von g. □

2.2.25 Folgerung: Stimmen die Werte der Polynome

$$f(x) = a_n x^n + \ldots + a_1 x + a_0$$
$$g(x) = b_n x^n + \ldots + b_1 x + b_0$$

an $n+1$-Stellen überein, so gilt $a_k = b_k$ für alle $k = 0, 1, \ldots, n$ und damit $f = g$.

In diesem Fall hat das Polynom $h\colon = f - g$ vom Grad $\leq n$ nämlich $n+1$ Nullstellen und muss daher das Nullpolynom sein.

Nullstellen von quadratischen und kubischen Gleichungen.

2.2.26 Satz: Ein quadratisches Polynom $f(z) = az^2 + bz + c \in \mathbb{C}[z]$ hat in $\mathbb{C}$ eine Nullstelle.

Beweis: Nach der *p-q-Formel*)sind die Nullstellen $\alpha = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$. $\square$

Wir wenden uns jetzt Polynomen vom Grad 3 zu:

$$f(z) = az^3 + bz^2 + cz + d \in \mathbb{C}[z]$$

Wir betrachten statt f das einfachere Polynom

$$p(z) = \frac{1}{a} f\left(z - \frac{b}{3a}\right) = z^3 + pz + q$$

mit geeigneten $p, q \in \mathbb{C}$. Es gilt:

α ist Nullstelle von $f \iff \alpha + \frac{b}{3a}$ ist Nullstelle von $p(z)$.

Sei $D := \left(\frac{p}{3}\right)^3 + \left(\frac{q}{2}\right)^2$ und

$$u_\pm := \sqrt[3]{-\frac{q}{2} \pm \sqrt{D}}.$$

Dann ist $\alpha = u_+ + u_-$ Nullstelle von $p(z)$ (ohne Beweis; dies ist die *Formel von Cardano*).

Beispiel: Sei $p(z)\colon = z^3 - 15z - 4$.
$\Rightarrow D = -125 + 4 = -121$
$\Rightarrow u_\pm = \sqrt[3]{+2 \pm 11i} = 2 \pm i \Rightarrow \alpha = 4$.
(denn $(2+i)^3 = (3+4i)(2+i) = 2 + 11i$).
$\Rightarrow \alpha = 4$ ist Nullstelle von $p(z)$.

$\Rightarrow p(z) = (z-4)q(z)$
mit $q(z) = az^2 + bz + c$ für geeignete $a, b, c \in \mathbb{C}$. Aus einer leichten Rechnung ergibt sich $q(z) = z^2 + 4z + 1$. Damit sind $-2 \pm \sqrt{3}$ die weiteren Nullstellen von $p(z)$ und es gilt:

$$p(z) = (z-4)(z+2+\sqrt{3}))(z+2-\sqrt{3})).$$

Man kann auch noch für Polynome vom Grad 4 die Nullstellen explizit durch seine Koeffizienten berechnen. Für Polynome vom Grad ≥ 5 ist das i.a. nicht mehr möglich, wie der norwegische Mathematiker N.H. Abel 1826 gezeigt hat.

Der folgende Satz garantiert die Existenz von Nullstellen von Polynomen über $\mathbb{C}$.

2.2.27 Satz (Fundamentalsatz der Algebra): Jedes nicht-konstante Polynom $f(z)$ (d.h. $\deg f > 0$) über $\mathbb{C}$ besitzt eine Nullstelle.

Ein Beweis findet sich z.B. in dem Lehrbuch Analysis 1 von K. Königsberger.

2.2.28 Folgerung (Satz von der Linearfaktorzerlegung): Jedes nicht konstante Polynom $f \in \mathbb{C}[z]$ besitzt eine Darstellung

$$f(z) = a(z-\alpha_1) \cdot \ldots \cdot (z-\alpha_n)$$

Dabei sind die Faktoren $(z-\alpha_1), \ldots, (z-\alpha_n)$ bis auf die Reihenfolge eindeutig bestimmt.

Beweis: Ist α_1 eine Nullstelle von f so gilt: $f(z) = (z-\alpha_1)f_1$ wobei f_1 eine Polynom vom Grad $\deg f - 1$ ist. Ist $\deg f_1 > 0$ so schreiben wir wieder $f_1 = (z-\alpha_2)f_2$ wobei α_2 Nullstelle von f_1 ist, also $f = (z-\alpha_1)(z-\alpha_2)f_2$. So fortfahrend erhält man schliesslich die gesuchte Darstellung. $\square$

3 Vektorräume und lineare Abbildungen

3.1 Vektorräume

Sei K ein Körper (etwa $K = \mathbb{R}$ oder $= \mathbb{C}$).

3.1.1 Definition: Ein K-Vektorraum oder ein Vektorraum über K ist eine Menge V zusammen mit zwei Abbildungen

$$+ : V \times V \to V, (v, w) \mapsto v + w \qquad \text{(Addition)}$$

und

$$\cdot : K \times V \to V, (\lambda, v) \mapsto \lambda v \qquad \text{(skalare Multiplikation)}$$

so dass gilt:

(1) $(V, +)$ ist eine abelsche Gruppe.

(2) $\lambda \cdot (\mu \cdot v) = (\lambda\mu) \cdot v \qquad \forall\, \lambda, \mu \in K, v \in V$.

(3) $1 \cdot v = v \qquad \forall\, v \in V$.

(4) $\lambda(v + w) = \lambda v + \lambda w$ und $(\lambda + \mu)v = \lambda v + \mu v$ für alle $\lambda, \mu \in K$ und $v, w \in V$.

Die Elemente in V werden als Vektoren bezeichnet. Mit 0_V (oder einfach 0) bezeichnen wir das neutrale Element von $(V, +)$ (es wird der Nullvektor genannt). Das Inverse von $v \in V$ in $(V, +)$ bezeichnen wir wie üblich mit $-v$. Für $v, w \in V$ setzen wir $v - w :\, = v + (-w)$

Ist $K = \mathbb{R}$ (bzw. $= \mathbb{C}$), so spricht man auch von einem reellen (bzw. komplexen) Vektorraum.

3.1.2 Bemerkung: Sei V ein K-Vektorraum. Dann gilt:

(a) $0 \cdot v = 0_V \qquad \forall\, v \in V$

(b) $\lambda \cdot 0_V = 0_V \qquad \forall\, \lambda \in K$

(c) $(-1) \cdot v = -v \qquad \forall\, v \in V$.

Beweis: Es gilt

$$\begin{aligned} 0 \cdot v = 0 \cdot v + 0_V = 0 \cdot v + (0 \cdot v - (0 \cdot v)) = \\ = (0 \cdot v + 0 \cdot v) - (0 \cdot v) \overset{(4)}{=} (0 + 0) \cdot v - (0 \cdot v) = 0_V \end{aligned}$$

und damit (a). Der Beweis von (b) verläuft ganz analog. (c) ergibt sich wie folgt:

$$\begin{aligned} -v &\overset{(a)}{=} -v + (1 + (-1))v = -v + 1v + (-1)v = \\ &= -v + v + (-1)v = 0_V + (-1)v = (-1)v. \end{aligned}$$

$\square$

3.1.3 Beispiele: (a) Das Standardbeispiel für einen Vektorraum ist die Menge der Spaltenvektoren $K^n = M(n \times 1, K) = \{ \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \mid x_1, \dots, x_n \in K\}$ mit der Addition und skalaren Multiplikation

$$\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} + \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} x_1 + y_1 \\ \vdots \\ x_n + y_n \end{pmatrix}$$

$$\lambda \cdot \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} \lambda x_1 \\ \vdots \\ \lambda x_n \end{pmatrix}.$$

(b) Die Menge der Polynome $K[x]$ über K ist ein Vektorraum. Dabei ist die Addition und skalare Multiplikation definiert durch

$$f + g : K \to K, x \mapsto (f+g)(x) := f(x) + g(x) \tag{15}$$

$$\lambda f : K \to K, x \mapsto (\lambda \cdot f)(x) := \lambda f(x) \tag{16}$$

(c) Sei M eine Menge und K ein Körper. Die Menge $\operatorname{Abb}(M, K)$ der Abbildungen $f : M \to K$ ist ein K-Vektorraum. Dabei ist die Addition und Skalramultiplikation wieder durch (16) definiert, d.h. es gilt $(f+g)(m) := f(m) + g(m)$ und $(\lambda f)(m) := \lambda f(m)$ für alle $f, g \in \operatorname{Abb}(M, K)$ und $\lambda \in K$.

3.1.4 Definition: Sei V ein K-Vektorraum. Eine Teilmenge U von V heisst Untervektorraum, wenn $U \neq \emptyset$ und für alle $v, w \in U$, und alle $\lambda \in K$ gilt:

$$v + w, \quad \lambda v \in U.$$

3.1.5 Beispiele: (a) Die Untervektorräume von $\mathbb{R}^2$ sind $\{0\}$, $\mathbb{R}^2$ und die Geraden durch den Nullpunkt $(0,0)$, d.h. Mengen der Form $\{(x,y) \in \mathbb{R}^2 \mid ax + by = 0\}$ wobei a und b reelle Zahlen sind, die nicht beide $= 0$ sind.

(b) Sei K ein Körper und

$$\begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = 0 \tag{$*$}$$

ein homogenes lineares Gleichungssystem (über K). Dann ist die Lösungsmenge U von $(*)$, d.h. die Menge der n-Tupel reeller Zahlen $\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in K^n$ die $(*)$ erfüllen ein Untervektorraum von K^n.

(c) Sei V der $\mathbb{R}$-Vektorraum der Funktionen $f : \mathbb{R} \to \mathbb{R}$. Sei

$$U := \{f \in V \mid f \text{ ist } 2\times \text{ stetig differenzierbar und } f'' + f = 0\}.$$

Dann ist U ein Untervektorraum von V. Zum Beispiel sind $\sin(x)$ und $\cos(x)$ Elemente von U, da $\sin' = \cos$ und $\cos' = -\sin$.

3.1.6 Bemerkung: Sei V ein K-Vektorraum und $U \subseteq V$ ein Untervektorraum. Dann ist U selbst ein Vektorraum (genauer gilt: U zusammen mit der Addition $+\big|_{U\times U}$ und skalaren Multiplikation $\cdot\big|_{K\times U}$) ist ein K-Vektorraum).

Beweis: Alle Axiome bis auf (1) folgen sofort aus der Tatsache, dass sie für V gelten.

Für $u \in U$ gilt

$$-u = (-1) \cdot u \in U.$$

Da $U \neq \emptyset$ ist U eine Untergruppe von $(V, +)$. Damit gilt (1) auch für U. $\square$

3.1.7 Bemerkung: Sind U_1, U_2 Untervektorräume von V, so ist auch $U_1 \cap U_2$ ein Untervektorraum.

Als nächstes diskutieren wir die Begriffe *Linearkombinationen, lineare Unabhängigkeit, Erzeugendensystem* und *Basis.*

Sei V ein K-Vektorraum. Seien $v_1, \ldots, v_n$ Vektoren von V.

3.1.8 Definition: Ein Vektor der Form

$$\lambda_1 v_1 + \ldots + \lambda_n v_n$$

mit $\lambda_1, \ldots, \lambda_n \in K$, heisst Linearkombination von $(v_1, \ldots, v_n)$. Die Menge aller Linearkombinationen von $(v_1, \ldots, v_n)$ wird mit $L(v_1, \ldots, v_n)$ bezeichnet. Also

$$L(v_1, \ldots, v_n) = \{\lambda_1 v_1 + \ldots + \lambda_n v_n \mid \lambda_1, \ldots, \lambda_n \in K\}.$$

Die Menge $L(v_1, \ldots, v_n)$ heisst die lineare Hülle von $(v_1, \ldots, v_n)$ (oder auch der von den Vektoren $v_1, \ldots, v_n$ aufgespannten oder erzeugten Raum).

3.1.9 Bemerkung: (a) $L(v_1, \ldots, v_n)$ ist ein Untervektorraum von V.

(b) Ist $U \subseteq V$ ein Untervektorraum und $v_1, \ldots, v_n \in U$ so gilt $L(v_1, \ldots, v_n) \subseteq U$. $L(v_1, \ldots, v_n)$ ist also der kleinste Untervektorraum der die Vektoren $v_1, \ldots, v_n$ enthält.

Beweis von (a): Die Summe zweier Linearkombinationen und skalare Vielfache einer Linearkombination ist wieder eine Linearkombination, denn

$$(\lambda_1 v_1 + \ldots + \lambda_n v_n) + (\mu_1 v_1 + \ldots + \mu_n v_n) = (\lambda_1 + \mu_1)v_1 + \ldots + (\lambda_n + \mu_n)v_n$$
$$\lambda(\lambda_1 v_1 + \ldots + \lambda_n v_n) = (\lambda\lambda_1)v_1 + \ldots + (\lambda\lambda_n)v_n.$$

□

3.1.10 Beispiele: (a) Sei $v = \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2$ ein Vektor $\neq 0$. Dann ist $L(v)$ die Gerade in $\mathbb{R}^2$ durch die Punkte $\begin{pmatrix} 0 \\ 0 \end{pmatrix}$ und $\begin{pmatrix} x \\ y \end{pmatrix}$.

(b) Seien $v_1 := \begin{pmatrix} 1 \\ 2 \end{pmatrix}, v_2 := \begin{pmatrix} 2 \\ 5 \end{pmatrix}, w := \begin{pmatrix} 2 \\ 3 \end{pmatrix} \in \mathbb{R}^2$. Um zu entscheiden, ob w Linearkombination von (v_1, v_2) ist müssen wir das lineare Gleichungssystem

$$\begin{array}{rcrcr} x & + & 2y & = & 2 \\ 2x & + & 5y & = & 3 \end{array}$$

untersuchen. Offenbar gibt es eine eindeutig bestimmte Lösung $(x, y) = (-1, 4)$. Also, $w = (-1)v_1 + 4v_2$.

(c) Das Polynom $x^3 \in \mathrm{Abb}\,(\mathbb{R}, \mathbb{R})$ ist Linearkombination der Polynome $(x-1)^k$, $k = 0, 1, 2, 3$. Es gilt nämlich

$$x^3 = ((x-1)+1)^3 = 1 \cdot (x-1)^3 + 3(x-1)^2 + 3(x-1) + 1 \cdot (x-1)^0$$

Umgekehrt ist die Funktion $(1+x)^n$ eine Linearkombination der *Monome* $1, x, x^2, \ldots, x^n$, denn

$$(1+x)^n = \binom{n}{0} 1 + \binom{n}{1} x^1 + \ldots + \binom{n}{n-1} x^{n-1} + \binom{n}{n} x^n.$$

3.1.11 Definition: Sei V ein K-Vektorraum.

(a) Ein n-Tupel $(v_1, \ldots, v_n)$ von Vektoren von V heisst Erzeugendensystem, wenn jeder Vektor von V Linearkombination von $v_1, \ldots, v_n$ ist, d.h. wenn $V = L(v_1, \ldots, v_n)$.

(b) V heisst endlich erzeugt, wenn ein Erzeugendensystem endlicher Länge existiert, d.h. wenn es eine natürliche Zahl n und ein Erzeugendensystem $(v_1, \ldots, v_n)$ von V gibt.

3.1.12 Beispiele: (a) Die Vektoren $v_1 \colon = \begin{pmatrix} 1 \\ 2 \end{pmatrix}, v_2 \colon = \begin{pmatrix} 2 \\ 5 \end{pmatrix}$ bilden ein Erzeugendensystem von $\mathbb{R}^2$.

(b) $(e_1, \ldots, e_n)$ ein Erzeugendensystem von K^n. Insbesondere ist K^n endlich erzeugt.

(c) Ist K unendlich so ist $K[x]$ nicht endlich erzeugt.

(Beweis: Angenommen $K[x]$ wäre endlich erzeugt. Sei $(f_1, \ldots, f_n)$ ein Erzeugendensystem. Das Maximum der Grade von $f_1, \ldots, f_n$ sei N. Wir schreibe das Polynom z^{N+1} als Linearkombination von $(f_1, \ldots, f_n)$

$$z^{N+1} = \lambda_1 f_1 + \ldots + \lambda_n f_n$$

Auf der linken Seite der Gleichung steht ein Polynom vom Grad $N + 1$ während das Polynom rechts vom Grad $\leq N$ ist. Widerspruch!)

Die Aufgabe einen (Unter-)Vektorraum durch möglichst wenige Vektoren zu erzeugen führt zum Begriff der *linearen Unabhängigkeit*.

3.1.13 Definition: Sei V ein K-Vektorraum. Ein n-Tupel von Vektoren $(v_1, \ldots, v_n)$ heißt linear unabhängig, falls für $\lambda_1, \ldots, \lambda_n \in K$ gilt:

$$\lambda_1 v_1 + \ldots + \lambda_n v_n = 0 \Rightarrow \lambda_1 = \ldots = \lambda_k = 0.$$

Das Tupel $(v_1, \ldots, v_n)$ heisst linear abhängig, falls es nicht linear unabhängig ist.

Anders ausgedrückt: $(v_1, \ldots, v_k)$ ist linear abhängig, falls es eine Darstellung des Nullvektors als Linearkombination $0 = \lambda_1 \cdot v_1 + \ldots + \lambda_k \cdot v_k$ gibt, wobei die $\lambda_1, \ldots, \lambda_k \in K$ nicht alle gleich Nulll sind.

3.1.14 Beispiele: (a) Die Vektoren $x = (1, 0, 1)$, $y = (-1, 2, 1)$, $z = (0, 2, 2)$ sind linear abhängig, denn

$$3x + 3y - 3z = 0$$

(b) Die Vektoren $e_1, \ldots, e_n$ von K^n sind linear unabhängig. Aus

$$\lambda_1 \cdot e_1 + \ldots + \lambda_n \cdot e_n = 0$$

folgt nämlich $(\lambda_1, \ldots, \lambda_n) = (0, \ldots, 0)$, also $\lambda_1 = \ldots = \lambda_n = 0$.

(c) Ein einzelner Vektor v ist genau dann linear unabhängig, wenn $v \neq 0$.

(d) Zwei Vektoren v_1, v_2 sind linear unabhängig, wenn keiner Vielfaches des anderen ist.

(e) Ist K unendlich so sind die Monome $1 = x^0, x, x^2, \ldots, x^n$ in $K[x]$ linear unabhängig. Dies ergibt sich aus Folgerung 2.2.25.

(f) Drei Vektoren in $\mathbb{R}^3$ sind linear abhängig, wenn sie in einer Ebene liegen.

3.1.15 Lemma (Kriterien für lineare Unabhängigkeit): Sei V ein K-Vektorraum und $(v_1, \ldots, v_n)$ ein n-Tupel von Vektoren. Die folgenden Bedingungen sind äquivalent:

(i) $(v_1, \ldots, v_n)$ sind linear unabhängig.

(ii) Jeder Vektor in $L(v_1, \ldots, v_n)$ läßt sich in eindeutiger Weise als Linearkombination der $v_1, \ldots, v_n$ schreiben.

(iii) Keiner der Vektoren $v_1, \ldots, v_n$ ist Linearkombination der anderen.

Beweis: (i) $\Rightarrow$ (ii): Sei $v \in L(v_1, \ldots, v_n)$ und seinen

$$v = \lambda_1 v_1 + \ldots + \lambda_n v_n = \mu_1 v_1 + \ldots + \mu_n v_n$$

zwei Darstellungen von v als Linearkombination von $(v_1, \ldots, v_n)$.

$$\begin{aligned} &\Rightarrow && 0 = v - v = (\lambda_1 - \mu_1)v_1 + \ldots + (\lambda_n - \mu_n)v_n \\ &\stackrel{(i)}{\Rightarrow} && \lambda_1 - \mu_1 = \ldots = \lambda_n - \mu_n = 0 \\ &\Rightarrow && \lambda_1 = \mu_1, \ldots, \lambda_n = \mu_n. \end{aligned}$$

(i) $\Rightarrow$ (iii): Angenommen es existiert ein Index $i \in \{1, 2, \ldots, k\}$, so dass v_i Linearkombination von $(v_1, \ldots, v_{i-1}, v_{i+1}, \ldots, v_n)$ ist. Sei $v_i = \lambda_1 v_1 + \ldots + \lambda_{i-1}v_{i-1} + \lambda_{i+1}v_{i+1} + \ldots + \lambda_n v_n$ eine solche. Dann erhält man eine nicht-triviale Darstellung des Nullvektors als Linearkombination von $(v_1, \ldots, v_n)$ nämlich

$$0 = \lambda_1 v_1 + \ldots + \lambda_{i-1}v_{i-1} + (-1)v_i + \lambda_{i+1}v_{i+1} + \ldots + \lambda_k v_k$$

im Widerspruch zur linearen Unabhängigkeit von $(v_1, \ldots, v_n)$.

(iii) $\Rightarrow$ (i): Seien $\lambda_1, \ldots, \lambda_n \in K$ mit

$$\lambda_1 v_1 + \ldots + \lambda_n v_n = 0.$$

Angenommen $\lambda_i \neq 0$ für einen Index $i \in \{1, \ldots, n\}$. Dann folgt

$$v_i = -\frac{\lambda_1}{\lambda_i} v_1 - \ldots - \frac{\lambda_{i-1}}{\lambda_i} v_{i-1} - \frac{\lambda_{i+1}}{\lambda_i} v_{i+1} - \ldots - \frac{\lambda_n}{\lambda_i} v_n$$

im Widerspruch zur Annahme (iii). $\square$

3.1.16 Lemma: Es seien $k+1$ Vektoren $v_1, \ldots, v_k, v_{k+1}$ in V gegeben. Sind $(v_1, \ldots, v_k)$ linear unabhängig und ist $v_{k+1} \in V \setminus L(v_1, \ldots, v_k)$ so ist $(v_1, \ldots, v_k, v_{k+1})$ linear unabhängig.

Beweis: Seien $\lambda_1, \ldots, \lambda_{k+1} \in K$ mit

$$\lambda_1 v_1 + \ldots + \lambda_{k+1} v_{k+1} = 0. \tag{17}$$

Dann ist zunächst $\lambda_{k+1} = 0$, denn wäre $\lambda_{k+1} \neq 0$, so könnte man die Gleichung nach v_{k+1} auflösen

$$v_{k+1} = -\frac{\lambda_1}{\lambda_{k+1}} v_1 - \ldots - \frac{\lambda_k}{\lambda_{k+1}} v_k$$

und v_{k+1} wäre Linearkombination von $(v_1, \ldots, v_k)$ im Widerspruch zur Voraussetzung.

Also können wir den Summanden $\lambda_{k+1} v_{k+1}$ in (17) weglassen und erhalten

$$\lambda_1 v_1 + \ldots + \lambda_k v_k = 0.$$

Da $(v_1, \ldots, v_k)$ nach Voraussetzung linear unabhänig ist, folgt $\lambda_1 = \ldots = \lambda_k = 0$. $\square$

3.1.17 Lemma (Schrankenlemma): Sei V ein K-Vektorraum und $n \in \mathbb{N}$. Dann sind die folgenden beiden Aussagen äquivalent:

(i) V wird von n Vektoren erzeugt.

(ii) Je $n+1$ Vektoren von V sind linear abhängig.

Beweis: Ist $V = \{0\}$ so ist nichts zu zeigen (beide Aussagen sind erfüllt). Wir nehmen also an, dass $V \neq \{0\}$.

(i) $\Rightarrow$ (ii): Seien $v_1, \ldots, v_{n+1} \in V$. Wir wollen zeigen, dass es $\lambda_1, \ldots, \lambda_{n+1} \in K$ gibt, die nicht alle $= 0$ sind, so dass

$$\lambda_1 v_1 + \ldots + \lambda_{n+1} v_{n+1} = 0.$$

Nach Voraussetzung gibt es ein Erzeugendensystem $(b_1, \ldots, b_n)$ von V und damit $a_{ij} \in K, i = 1, \ldots, n, j = 1, \ldots, n+1$, so dass

$$v_j = a_{1j} b_1 + \ldots + a_{nj} b_n = \sum_{i=1}^{n} a_{ij} b_i \qquad \text{für } j = 1, \ldots, n+1.$$

Damit ist

$$\sum_{j=1}^{n+1} \lambda_j v_j = \sum_{j=1}^{n+1} \lambda_j \left(\sum_{i=1}^{n} a_{ij} b_i \right) = \sum_{i=1}^{n} \left(\sum_{j=1}^{n+1} a_{ij} \lambda_j \right) b_i.$$

Es genügt also zu zeigen, dass es Werte λ_j gibt, die nicht alle $= 0$ sind, so dass

$$\sum_{j=1}^{n+1} a_{ij} \lambda_j \qquad \text{für alle } i = 1, \ldots, n.$$

Das ist ein homogenes lineares Gleichungssystem mit n Gleichungen für $n+1$ Unbekannte. Es hat also immer eine nicht-triviale Lösung.

(ii) $\Rightarrow$ (i): Wir wählen uns ein linear unabhängiges Tupel $(b_1, \ldots b_r)$ von Vektoren von grösstmöglicher Länge r. Das existiert nach Voraussetzung und es ist $1 \leq r \leq n$. Sei $U := L(b_1, \ldots, b_r)$. Ist $U = V$, so sind wir fertig. Andernfalls gibt es ein $b_{r+1} \in V \smallsetminus U$. Nach dem obigen Lemma ist dann $(b_1, \ldots b_r, b_{r+1})$ linear unabhängig. Das ist ein Widerspruch zur Wahl von r. $\square$

3.1.18 Definition: Sei V ein K-Vektorraum. Ein n-Tupel $(b_1, \ldots, b_n)$ von Vektoren von V heisst Basis, wenn $(b_1, \ldots, b_n)$ ein Erzeugendensystem und linear unabhängig ist.

3.1.19 Bemerkung: Ist $(b_1, \ldots, b_n)$ ein Erzeugendensystem von V, so lässt sich jeder Vektor $v \in V$ schreiben als

$$v = \lambda_1 v_1 + \ldots + \lambda_n b_n$$

für gewisse $\lambda_1, \ldots, \lambda_n \in K$. Ist $(b_1, \ldots, b_n)$ eine Basis, so sind $\lambda_1, \ldots, \lambda_n$ nach dem Kriterium (ii) von Lemma 3.1.15 eindeutig bestimmt.

3.1.20 Beispiel: (a) $(e_1, \ldots, e_n)$ ist eine Basis von K^n. Sie heisst Standardbasis des K^n.

(b) Sei V ein reeller Vektorraum mit Basis (v_1, v_2). Dann ist auch (v_1+v_2, v_1-v_2) eine Basis.

3.1.21 Satz (Basissatz): Sei V ein endlich erzeugter K-Vektorraum. Dann gilt:

(a) V besitzt eine Basis. Genauer: Ist $(v_1, \ldots, v_m)$ ein Erzeugendensystem, so gibt es eine Basis $(b_1, \ldots, b_n)$ mit $b_i \in \{v_1, \ldots, v_m\}$ für alle $i \in \{1, \ldots, n\}$, d.h. man erhält eine Basis aus $(v_1, \ldots, v_m)$ indem man gegebenenfalls einige der Vektoren weglässt.

(b) Je zwei Basen haben die gleiche Länge.

(c) (Ergänzungssatz) Sind $b_1, \ldots, b_r$ linear unabhängig und keine Basis von V, so gibt es Vektoren $b_{r+1}, \ldots, b_n \in V$, so dass $(b_1, \ldots, b_n)$ eine Basis von V ist.

Beweis: Zu (a): Sei $(b_1, \ldots, b_n)$ ein Erzeugendensystem von V mit minimalem n, gebildet aus Vektoren in $\{v_1, \ldots, v_m\}$. Wir wollen zeigen, dass $(b_1, \ldots, b_n)$ linear unabhängig ist und damit eine Basis. Nach Lemma 3.1.15 genügt es zu zeigen, dass keiner der Vektoren $b_1, \ldots, b_n$ Linearkombination der anderen ist.

Wäre z.B. b_n Linearkombination von $(b_1, \ldots, b_{n-1})$, also $b_n \in L(b_1, \ldots, b_{n-1})$, so würde für eine beliebige Linearkombination $\sum_{i=1}^n \lambda_i b_i$ von $b_1, \ldots, b_n$ gelten

$$\sum_{i=1}^n \lambda_i b_i = (\lambda_1 b_1 + \ldots + \lambda_{n-1} b_{n-1}) + \lambda_n b_n \in L(b_1, \ldots, b_{n-1})$$

da $L(b_1, \ldots, b_{n-1})$ ein Untervektorraum ist. Also würde folgen

$$V = L(b_1, \ldots, b_n) \subseteq L(b_1, \ldots, b_{n-1}),$$

d.h. $(b_1, \ldots, b_{n-1})$ wäre ein Erzeugendensystem. Das steht im Widerspruch zur Minimalität von n.

Zu (b): Seien $(b_1, \ldots, b_n)$ und $(b'_1, \ldots, b'_{n'})$ zwei Basen von V. Wir wollen zeigen, dass $n = n'$. Da $(b_1, \ldots, b_n)$ ein Erzeugendensystem ist, ist $n' \leq n$, den das n'-Tupel $(b'_1, \ldots, b'_{n'})$ wäre sonst nach dem Schrankenlemma linear abhängig. Analog, indem man die Rollen von $(b_1, \ldots, b_n)$ und $(b'_1, \ldots, b'_{n'})$ vertauscht, ergibt sich $n \leq n'$, also $n = n'$.

Zu (c): V sei erzeugt von n Elementen. Sind $b_1, \ldots, b_r$ linear unabhängig aber keine Basis, so sind sie kein Erzeugendensystem, d.h. $L(b_1, \ldots, b_r) \neq V$. Dann gibt es ein $b_{r+1} \in V \setminus L(b_1, \ldots, b_r)$. Nach einem der vorangegangenen Lemmata ist $(b_1, \ldots, b_r, b_{r+1})$ linear unabhängig. Diesen Prozess können wir fortsetzen, solange wir nicht bei einem Erzeugendensystem angekommen sind. Aber nach dem Schrankenlemma geht das allenfalls, solange $r < n$ ist. Also bricht der Prozess ab, und wir erhalten eine Basis der gesuchten Form. $\square$

3.1.22 Definition: Sei V ein K-Vektorraum. Die Dimension von V

$$\dim(V) \in \{0, 1, 2, \ldots\} \cup \{\infty\}$$

ist wie folgt definiert:

1. Ist $V = \{0\}$, so sei $\dim V = 0$.
2. Ist $V \neq \{0\}$ endlich erzeugt und $(b_1, \ldots, b_n)$ eine Basis von V, so sei
$$\dim V = n.$$
3. Ist V nicht endlich erzeugt, so sei
$$\dim V = \infty.$$

Beispielsweise ist $\dim(K^n) = n$. Für den reellen Vektorraum $\mathbb{R}[x]$ aller Polynome über $\mathbb{R}$ gilt $\dim(\mathbb{R}[x]) = \infty$ und $\dim\{f(x) \in \mathbb{R}[x] \mid \deg f(x) \leq n\}) = n + 1$.

Ein Vektorraum V heisst endlich-dimensional falls $\dim(V) < \infty$, also wenn V endlich erzeugt ist. Im Folgenden betrachten wir haupsächlich endlich-dimensionale Vektorräume. Als Folgerung aus dem Basissatz erhalten wir:

3.1.23 Folgerung: Sei V ein endlich-dimensionaler Vektorraum der Dimension n und $(v_1, \ldots, v_n)$ ein n-Tupel von Vektoren V. Dann sind die folgenden Bedingungen äquivalent:

(i) $(v_1, \ldots, v_n)$ ist eine Basis.

(ii) $(v_1, \ldots, v_n)$ ist linear unabhängig.

(iii) $(v_1, \ldots, v_n)$ ist ein Erzeugendensystem.

Beweis: (ii) $\Rightarrow$ (i): Wäre $(v_1, \dots, v_n)$ keine Basis, so könnte man es nach dem Ergänzungssatz zu einer Basis auffüllen, die dann aber mehrals n Elemente enthielte im Widerspruch zu $\dim(V) = n$.

(iii) $\Rightarrow$ (i): Wäre $(v_1, \dots, v_n)$ keine Basis, so würde aus $(v_1, \dots, v_n)$ durch Weglassen einiger Vektoren eine Basis mit weniger als n Elementen entstehen. Das ist wieder ein Widerspruch zu $\dim(V) = n$. $\square$

3.1.24 Folgerung: (Dimension von Unterräumen) Sei $U \subseteq V$ ein Untervektorraum des endlich-dimensionalen Vektorraums V. Dann ist auch U endlich-dimensional und

$$\dim U \leq \dim V.$$

Gleichheit gilt genau dann, wenn $U = V$ ist.

Beweis: Sei $n\colon = \dim(V)$. Nach dem Schrankenlemma sind dann je $n + 1$ Vektoren in V linear abhängig. Also sind auch je $n+1$ Vektoren aus U linear abhängig. Nach dem Schrankenlemma (angewendet auf U) hat U daher ein Erzeugendensystem aus n Elementen und folglich, nach Teil (i) des Basissatz, eine Basis aus höchstens n Vektoren. Das zeigt, dass $\dim(U) \leq \dim(V)$.

Ist $\dim(U) = \dim(V) = n$ und $(v_1, \dots, v_n)$ eine Basis von U, so ist $(v_1, \dots, v_n)$, aufgrund der Implikation (ii) $\Rightarrow$ (i) in der obigen Folgerung 1, auch eine Basis von V und damit $U = L(v_1, \dots, v_n) = V$. $\square$

3.1.25 Definition: Seien U_1, U_2 Untervektorräume eines K-Vektorraums V. Die Menge

$$U_1 + U_2\colon = \{u_1 + u_2 \mid u_1 \in U_1, u_2 \in U_2\}$$

heisst die Summe von U_1 und U_2. Es ist der kleinste Untervektorraum von V der U_1 und U_2 enthält.

3.1.26 Folgerung: (Dimensionsformel für Unterräume) Sei V ein K-Vektorraum und seien U_1, U_2 endlich-dimensionale Untervektorräume. Dann ist auch $U_1 + U_2$ endlich-dimensional und es gilt

$$\dim(U_1 + U_2) + \dim(U_1 \cap U_2) = \dim(U_1) + \dim(U_2).$$