NOTIZEN ZUR VORLESUNG LINEARE ALGEBRA

WERNER BALLMANN

Dies ist ein vorläufiges Skript zur Vorlesung Lineare Algebra. Ich werde nicht mit der Vorlesung Schritt halten können. Es wird deshalb immer wieder Ergänzungen zum Skript geben, ausserdem werde ich immer wieder Korrekturen vornehmen müssen.

Ich bedanke mich bei Frau Pratoussevitch und den Herren Harrach, Kouris, Krüger, Rötzel und Schlippe für wertvolle Hinweise und Korrekturvorschläge.

1. WORUM GEHT ES?

Dazu werde ich vielleicht später etwas mehr schreiben. Zunächst sollten Sie an lineare Gleichungen denken. Es wird sich dabei aber nicht um zwei Gleichungen in zwei Unbekannten handeln, sondern um m Gleichungen in n Unbekannten. Sie dürfen sich m und n groß vorstellen. Was kann man über die Lösungsmengen solcher Gleichungssysteme sagen? Wann gibt es Lösungen? Gibt es mehr als eine Lösung?

Anfangen werde ich mit einem Verfahren aus der Kryptographie, dem sogenannten RSA-Verfahren. Auf dem Weg zu seiner Herleitung werden wir einige Begriffe kennenlernen, die grundlegend für die Lineare Algebra sind, insbesondere die Begriffe Gruppe, Ring und Körper. Einige der Resultate gehören zur elementaren Zahlentheorie. Ich hoffe, Sie haben Spass an und mit Zahlen.

2. WAS ICH VORAUSSETZE!

Ich nehme an, dass Sie in der Schule die üblichen Zahlbereiche kennengelernt haben. Zur Sicherheit und weil es in der Literatur geringfügige Differenzen in den Bezeichnungen gibt, zähle ich hier die Zahlbereiche zusammen mit den von mir verwendeten Bezeichnungen auf:

(1) die *natürlichen Zahlen* $\mathbb{N} = \{0, 1, 2, \ldots\}$,
(2) die *ganzen Zahlen* $\mathbb{Z} = \{\ldots, -2, -1, 0, 1, 2, \ldots\}$,
(3) die *rationalen Zahlen* $\mathbb{Q}$, also die Menge aller Brüche ganzer Zahlen,
(4) die *reellen Zahlen* $\mathbb{R}$, die "Zahlen auf der Zahlgeraden".

Die reellen Zahlen werden in der Vorlesung zur Analysis genauer untersucht. Vielleicht haben Sie in der Schule auch schon die Menge $\mathbb{C}$ der komplexen Zahlen kennengelernt, aber das setze ich nicht voraus. Ich werde die komplexen Zahlen später diskutieren.

Addition und Multiplikation von Zahlen aus den eben genannten Zahlbereichen sind Verknüpfungen, die zwei gegebenen Zahlen a und b ihre Summe bzw. ihr Produkt

Date: March 7, 2002.

zuordnen,

$$(a,b) \mapsto a+b \quad \text{bzw.} \quad (a,b) \mapsto a \cdot b.$$

Summe und Produkt genügen den Ihnen sicher bekannten Regeln,

$$\begin{aligned}
(a+b)+c &= a+(b+c) && \text{Assoziativgesetz der Addition,}\\
(a\cdot b)\cdot c &= a\cdot(b\cdot c) && \text{Assoziativgesetz der Multiplikation,}\\
a+b &= b+a && \text{Kommutativgesetz der Addition,}\\
a\cdot b &= b\cdot a && \text{Kommutativgesetz der Multiplikation,}\\
a\cdot(b+c) &= a\cdot b + a\cdot c && \text{Distributivgesetz.}
\end{aligned}$$

Wie üblich schreibe ich das Produkt zweier Zahlen einfach als ab anstelle von $a \cdot b$. Ferner benütze ich die Ihnen sicher auch geläufige Konvention $a - b := a + (-b)$. Der Doppelpunkt vor dem Gleichheitszeichen deutet an, dass ich die linke Seite der Gleichung durch die rechte Seite definiere. Ferner halte ich mich an die Konvention "Punkt geht vor Strich".

Die reellen Zahlen stellen wir uns nach der Größe angeordnet vor: Für $a, b \in \mathbb{R}$ schreibe ich $a < b$ wenn $b - a$ positiv ist. Analog benütze ich die Symbole $>, \geq$ und $\leq$. Ich nehme an, dass Sie mit dieser Schreibweise schon vertraut sind. Für $a \in \mathbb{R}$ ist der *Betrag* $|a|$ von a gegeben durch $|a| := a$ falls $a \geq 0$ bzw. $|a| := -a$ falls $a \leq 0$. Für $a, b, c \in \mathbb{R}$ gilt:

(1) $|a| \geq 0$; sogar $|a| > 0$, falls $a \neq 0$,
(2) $|ab| = |a| \cdot |b|$,
(3) $|a + b| \leq |a| + |b|$ (Dreiecksungleichung).

Wir werden an einigen Stellen benützen, dass der Betrag einer ganzen Zahl ≥ 1 ist, sofern die Zahl $\neq 0$ ist.

Mit Hilfe des Betrags wird der *Abstand* reeller Zahlen definiert: der Abstand von a zu b ist nach Definition $|a - b|$. Der Begriff des Abstands spielt in der Analysis eine sehr wichtige Rolle.

Eines der wichtigen Beweisprinzipien der Mathematik ist die vollständige Induktion. Eine prägnante Formulierung dieses Prinzips nenne ich

Prinzip des kleinsten Sünders [1]: Eine nichtleere Menge natürlicher Zahlen hat ein kleinstes Element.

[1] Für diesen Namen beanspruche ich keine Priorität.

3. Rechnen mit ganzen Zahlen

Wir befassen uns nun zunächst mit den ganzen Zahlen. Quelle und Referenz für diesen Abschnitt sind Kapitel 1 und der erste Abschnitt aus Kapitel 2 in [RU].

3.1. Satz (Kürzungsregel). *Seien $a, b, c \in \mathbb{Z}$ mit $a \neq 0$ und $ab = ac$. Dann ist $b = c$.*

Beweis. Wegen $ab = ac$ ist $a(b-c) = ab - ac = 0$. Wegen $a \neq 0$ folgt $b - c = 0$, also $b = c$. □

Wir haben dabei benützt, dass ein Produkt ganzer Zahlen nur dann 0 sein kann, wenn mindestens einer der Faktoren 0 ist.

3.2. Satz (Teilen mit Rest). *Seien $a, b \in \mathbb{Z}$, $b \geq 1$. Dann gibt es eindeutig bestimmte ganze Zahlen q, r mit $0 \leq r < b$, so dass $a = qb + r$. Falls $a \geq 0$, so ist auch $q \geq 0$. Wir nennen q den* Quotienten *und r den* Rest *bei der Division von a durch b.*

Beweis. Zunächst beweisen wir die Existenz von q und r. Sei dazu

$$A = \{n \in \mathbb{N} \mid \text{es gibt } k \in \mathbb{Z} \text{ mit } a - kb = n\}.$$

Nach Definition ist $A \subset \mathbb{N}$.

Falls $a \geq 0$ ist, so ist $a - 0 \cdot a = a \in \mathbb{N}$ und damit $a \in A$. Falls $a < 0$ ist, so ist $a - ab = a(1-b) \in A$, weil $1 - b \leq 0$ ist. Damit folgt $A \neq \emptyset$. Mit dem Prinzip des kleinsten Sünders folgt, dass A ein kleinstes Element hat. In weiser Voraussicht nennen wir dieses r. Wegen $A \subset \mathbb{N}$ ist $r \geq 0$.

Nach Definition von A gibt es eine ganze Zahl q mit $r = a - qb$. Falls nun $r \geq b$ wäre, so wäre

$$a - (q+1)b = a - qb - b = r - b \geq 0$$

und damit $a - (q+1)b \in A$. Wegen $b \geq 1$ ist aber $r - b < r$, damit hätten wir einen Widerspruch zur Wahl von r. Also ist $0 \leq r < b$ und $a = qb + r$ wie verlangt. Ferner ist klar, dass $q \geq 0$ ist, falls $a \geq 0$ ist.

Es bleibt, die Eindeutigkeit von q und r nachzuweisen. Seien dazu weitere ganze Zahlen q' und r' gegeben mit $0 \leq r' < b$ und $a = q'b + r'$. Nach Voraussetzung gilt $0 \leq r < b$ und $-b < -r' \leq 0$, also ist

$$(*) \qquad -b < r - r' < b \quad \text{bzw.} \quad |r - r'| < b.$$

Andererseits ist $(q' - q)b = q'b - qb = r - r'$, also $|q' - q| \cdot b = |r - r'|$. Falls nun $q \neq q'$ wäre, so wäre $|q' - q| \geq 1$, denn $q' - q$ ist eine ganze Zahl. Damit folgte aber $|r' - r| = |q' - q| \cdot b \geq b$, im Widerspruch zu $(*)$. Daher ist $q = q'$ und damit

$$qb + r = q'b + r' = qb + r',$$

also auch $r = r'$. □

3.3. Definition. Seien $a, b \in \mathbb{Z}$. Dann heisst b ein *Teiler* von a, wenn es ein $q \in \mathbb{Z}$ gibt mit $a = qb$. Wir schreiben dies als $b|a$.

Die 0 ist ein Sonderfall, alle ganzen Zahlen sind Teiler der 0. Offensichtlich gilt:

(1) $b|a \Leftrightarrow b|-a \Leftrightarrow -b|a$.
(2) $b|a \Leftrightarrow$ die Gleichung $a - bx = 0$ hat eine Lösung $x \in \mathbb{Z}$.

(3) $a = qb$ und $a \neq 0 \Rightarrow q \neq 0$ und $b \neq 0$.
(4) $a = qb$ und $b \neq 0 \Rightarrow a$ und b bestimmen q eindeutig.

Mit der Kürzungsregel folgt nämlich $q = q'$, falls $qb = q'b$.

3.4. RECHENREGELN. *Seien $a, b, c, d \in \mathbb{Z}$. Dann gilt:*

(1) $a|a$
(2) $a|b$ *und* $b|c \Rightarrow a|c$
(3) $a|b$ *und* $c|d \Rightarrow ac|bd$,
(4) $a|b$ *und* $a|c \Rightarrow a|(xb + yc)$ *für alle* $x, y \in \mathbb{Z}$.

Beweis. (1) $a = 1 \cdot a$.
(2) Falls $b = qa$ und $c = rb$, so ist $c = r(qa) = (rq)a$.
(3) Falls $b = qa$ und $d = rc$, so ist $bd = (qa)(rc) = (qr)(ac)$.
(4) Falls $b = qa$ und $c = ra$, so ist $xb + yc = (xq + yr)a$. □

3.5. SATZ. *Sei $a \in \mathbb{Z}$, $a \neq 0$, und $b \in \mathbb{Z}$ ein Teiler von a. Dann ist $1 \leq |b| \leq |a|$. Insbesondere hat a höchstens endlich viele verschiedene Teiler.*

Beweis. Sei q der Quotient, $a = qb$. Dann ist $|a| = |q| \cdot |b|$. Wegen $a \neq 0$ ist $q \neq 0$. Also ist $|q| \geq 1$, denn $q \in \mathbb{Z}$. Daher

$$|a| = |q| \cdot |b| \geq |b|.$$

Wegen $a \neq 0$ ist $b \neq 0$, es kommen also höchstens die Zahlen $\pm 1, \ldots, \pm|a|$ als Teiler von a in Frage. □

3.6. KOROLLAR. *Seien a, b positive ganze Zahlen. Falls dann $a|b$ und $b|a$, so ist $a = b$.*

Beweis. Nach Satz 3.5 ist $a \leq b$ und $b \leq a$, also $a = b$. □

In der Regel beschränkt man sich bei Teilbarkeitsfragen auf positive ganze Zahlen. Jede positive ganze Zahlen hat die 1 und sich selber als positive ganze Teiler. Die 1 ist die einzige positive ganze Zahl mit nur einem positiven Teiler.

3.7. DEFINITION. Eine positive ganze Zahl p heisst *Primzahl*, wenn p genau zwei positive Teiler hat.

Die Zahlen $2, 3, 5, 7, 11$ sind prim, aber $1, 4, 6, 8, 9, 10$ nicht.

3.8. SATZ. *Sei $p \in \mathbb{Z}$, $p > 1$. Dann sind äquivalent:*

(1) *p ist eine Primzahl.*
(2) *Für alle $a, b \in \mathbb{Z}$ gilt: $p|ab \Rightarrow p|a$ oder $p|b$.*

Beweis. (1) $\Rightarrow$ (2): Seien $a, b \in \mathbb{Z}$. Falls $a = 0$ oder $b = 0$, dann $p|a$ oder $p|b$, und wir sind fertig. Wir nehmen daher jetzt an, dass $a \neq 0$ und $b \neq 0$. Ohne Beschränkung der Allgemeinheit können wir auch annehmen, dass a und b positiv sind, also $a, b \geq 1$.

Sei nun $A = \{n \in \mathbb{N} \mid n \geq 1 \text{ und } p|an\}$. Dann sind $b, p \in A$, insbesondere $A \neq \emptyset$. Daher hat A ein kleinstes Element, k. Nach Definition von A ist $k \geq 1$.

Sei $n \in A$. Wir schreiben $n = qk + r$ mit $0 \leq r < k$. Wegen $p|an$, $p|ak$ und $ar = an - qak$ folgt $p|ar$. Nun ist $r \notin A$ wegen der Minimalität von k, daher $r = 0$. Damit folgt $k|n$ für jedes $n \in A$. Insbesondere folgt $k|p$, denn $p \in A$.

Nun ist p eine Primzahl. Es gibt daher nur zwei Fälle: $k = 1$ oder $k = p$. Im ersten Fall ist p Teiler von a. Im zweiten Fall folgt $p|b$, denn $b \in A$ und $k = p$ teilt jedes Element von A.

(1) $\Leftarrow$ (2): Sei $p = ab$ mit positiven ganzen Zahlen a, b. Nach Satz 3.5 ist $1 \leq a, b \leq p$. Mit (2) folgt $p|a$ oder $p|b$. Im ersten Fall folgt dann analog $p \leq a$, also $p = a$ und damit $b = 1$ nach der Kürzungsregel. Im zweiten Fall folgt $p \leq b$, also $p = b$ und damit $a = 1$. Insgesamt folgt, dass p nur 1 und p als positive Teiler hat. □

3.9. SATZ. *Sei a eine ganze Zahl > 1. Dann ist der kleinste Teiler $b > 1$ von a prim.*

Der kleinste Teiler $b > 1$ von a existiert, denn er ist ein kleinstes Element in der Menge $A = \{n \in \mathbb{N} \mid n \geq 2 \text{ und } n|a\}$ — diese Menge ist nicht leer, weil sie a enthält.

Beweis von Satz 3.9. Wäre b nicht prim, so gäbe es $k, l \geq 2$ mit $b = kl$. Aber dann wären $k, l < b$ und k und l Teiler von a im Widerspruch zur Wahl von b. □

3.10. SATZ (Euklid). [2] *Es gibt unendlich viele Primzahlen.*

Wir beweisen etwas mehr: Falls $p_1, \ldots, p_n$ Primzahlen sind, so ist keine dieser Zahlen Teiler der Zahl $a = p_1 \cdot \ldots \cdot p_n + 1$. Der kleinste Teiler > 1 von a ist damit eine von den gegebenen Primzahlen verschiedene Primzahl.

Beweis. Seien $p_1, \ldots, p_n$ Primzahlen und $a = p_1 \cdot \ldots \cdot p_n + 1$. Sei p der kleinste Teiler > 1 von a. Nach Satz 3.9 ist p prim. Falls $p = p_j$ für ein $j \in \{1, \ldots, n\}$, so wäre p auch ein Teiler von $p_1 \cdot \ldots \cdot p_n = a - 1$. Dann wäre p aber ein Teiler von $a - (a - 1) = 1$ im Widerspruch zu $p > 1$. □

Als Folgerung der Charakterisierung von Primzahlen weiter oben erhalten wir das

3.11. KOROLLAR. *Teilt eine Primzahl p ein Produkt $a_1 \cdot \ldots \cdot a_n$ ganzer Zahlen, dann teilt p einen der Faktoren.*

Beweis. Per Induktion über n: Der Fall $n = 1$ ist klar. Falls die Behauptung für $n - 1 \geq 1$ Faktoren richtig ist und p das Produkt $a_1 \cdot \ldots \cdot a_n$ ganzer Zahlen teilt, so teilt p nach Satz 3.8 entweder $(a_1 \cdot \ldots \cdot a_{n-1})$ oder a_n. Im ersten Fall teilt p nach Induktionshypothese einen der Faktoren $a_1, \ldots, a_{n-1}$, im zweiten Fall den Faktor a_n. □

3.12. DEFINITION. Eine Primzahl, die eine ganze Zahl a teilt, heißt ein *Primteiler* oder *Primfaktor* von a. Eine Darstellung $a = p_1 \cdot \ldots \cdot p_n$ von a als Produkt von Primzahlen heißt eine *Primzerlegung* von a.

3.13. BEISPIELE. Einige einfache Beispiele: $6 = 2 \cdot 3$, $8 = 2 \cdot 2 \cdot 2$, $15 = 3 \cdot 5$.

3.14. KONVENTION. Das *leere Produkt* hat 0 Faktoren, wir geben ihm den Wert 1. Entsprechend hat die *leere Summe* 0 Summanden, wir erteilen ihr den Wert 0.

3.15. SATZ (Hauptsatz der elementaren Zahlentheorie). *Jede positive ganze Zahl hat eine Primzerlegung. Bis auf die Reihenfolge der Faktoren ist diese eindeutig.*

[2] Euklid lebte um 300 v.u.Z.

Beweis. Zunächst zeigen wir die Existenz der Zerlegung, und zwar per Induktion über die gegebene Zahl a: Für $a = 1$ gilt die Behauptung nach der gerade getroffenen Konvention. Sei nun $a > 1$ und die Behauptung richtig für alle ganzen Zahlen b mit $1 \leq b < a$. Sei p_1 der kleinste positive Teiler von a, der > 1 ist. Dann ist p_1 eine Primzahl und $a = p_1 b$ für eine ganze Zahl b. Weil $a, p_1 > 0$ sind, ist auch $b > 0$. Wegen $p_1 > 1$ folgt $a = p_1 b > 1 \cdot b = b$. Nach Induktionshypothese ist $b = p_2 \dots p_n$ mit Primzahlen $p_2, \dots, p_n$. Damit $a = p_1 p_2 \cdot \ldots \cdot p_n$.

Wir zeigen jetzt die Eindeutigkeit der Zerlegung, dieses Mal per Induktion nach der Anzahl n der Primfaktoren: Für $n = 0$ (und auch für $n = 1$) ist die Aussage offenbar richtig. Sei nun $n \geq 1$ und

$$a = p_1 \cdot \ldots \cdot p_n = q_1 \cdot \ldots \cdot q_m$$

mit Primzahlen $p_1, \dots, p_n, q_1, \dots, q_m$. Sei p der kleinste positive Teiler von a, der > 1 ist. Nach Satz 3.9 ist p prim und teilt daher nach Korollar 3.11 einen der Faktoren $p_1, \dots, p_n$ und einen der Faktoren $q_1, \dots, q_m$. Nach eventueller Umnummerierung folgt $p | p_n$ und $p | q_m$. Weil p_n und q_m prim sind folgt $p = p_n$, $p = q_m$. Mit der Kürzungsregel erhalten wir

$$b := p_1 \dots p_{n-1} = q_1 \dots q_{m-1}.$$

Mit der Induktionshypothese folgt $m = n$ und dass die Primfaktoren $p_1, \dots, p_{n-1}$ mit den Primfaktoren $q_1, \dots, q_{n-1}$ bis auf die Reihenfolge übereinstimmen. □

3.16. KONVENTION. Wir ordnen die Primfaktoren der Größe nach, gleiche Faktoren fassen wir zu Potenzen zusammen:

$$3500 = 2 \cdot 2 \cdot 5 \cdot 5 \cdot 5 \cdot 7 = 2^2 \cdot 5^3 \cdot 7.$$

3.17. KOROLLAR. *Seien a, b positive ganze Zahlen und*

$$a = p_1^{n_1} \cdot \ldots \cdot p_r^{n_r}, \quad r \geq 0 \text{ und } n_1, \dots, n_r \geq 1,$$

die Primzerlegung von a. Dann ist b genau dann ein Teiler von a, wenn b ein Produkt der Form $b = p_1^{m_1} \cdot \ldots \cdot p_r^{m_r}$ ist mit $0 \leq m_i \leq n_i$ für $1 \leq i \leq r$.

Beweis. Es ist klar, dass alle Zahlen der angegebenen Form Teiler von a sind. Sei umgekehrt b ein Teiler von a. Um Trivialitäten zu vermeiden sei $b \neq 1$ und $b \neq a$, also $a = bb'$ mit $b' > 1$. Seien $b = q_1 \cdot \ldots \cdot q_s$ und $b' = q_1' \cdot \ldots \cdot q_t'$ die Primzerlegungen von b und b'. Dann ist

$$a = bb' = q_1 \cdot \ldots \cdot q_s \cdot q_1' \cdot \ldots \cdot q_t'$$

eine Primzerlegung von a. Wegen der Eindeutigkeit dieser ist b daher von der angegebenen Form. □

Sei a wie in Satz 3.15 und $\tau(a)$ die Anzahl der positiven Teiler von a. Spielt man die verschiedenen Möglichkeiten für die positiven Teiler von a durch, dann erhält man ohne grössere Probleme die Formel $\tau(a) = \Pi_{i=1}^{r}(n_i + 1)$.

Sei $\sigma(a)$ die Summe aller positiven Teiler von a. Dann heißt a *vollkommen*, wenn $\sigma(a) = 2a$ ist. Beispiele sind

$$6 = 1 + 2 + 3,\ 28 = 1 + 2 + 4 + 7 + 14,\ 496.$$

Gerade vollkommene Zahlen sind genau die Zahlen von der Form $2^{p-1}(2^p - 1)$ mit $2^p - 1$ prim. Solche Primzahlen heissen *Mersennesche Primzahlen* [3]. Damit $2^p - 1$ prim ist, muss p selber auch prim sein. Das ist aber nicht hinreichend, siehe Abschnitt 3 aus Kapitel 1 in [RU]. Beispiele ungerader vollkommener Zahlen sind nicht bekannt.

3.18. DEFINITION. Seien $a, b \in \mathbb{Z}$. Wir nennen $d \in \mathbb{Z}$ einen *größten gemeinsamen Teiler* von a und b und schreiben $d = \mathrm{ggT}(a, b)$, wenn

(1) $d \geq 0$, $d|a$ und $d|b$.
(2) Falls $c \in \mathbb{Z}$ und $c|a$, $c|b$, so $c|d$.

Falls d und d' natürliche Zahlen sind mit $d|d'$ und $d'|d$, so ist $d = d'$. Also ist der $\mathrm{ggT}(a, b)$ von a und b eindeutig. Zu klären ist die Existenz. Für alle $a \in \mathbb{Z}$ ist $\mathrm{ggT}(a, 0) = |a|$. Wegen $\mathrm{ggT}(a, b) = \mathrm{ggT}(b, a) = \mathrm{ggT}(-a, b)$ bleibt uns, den Fall $a, b \geq 1$ zu untersuchen.

3.19. SATZ (Euklidischer Algorithmus). *Seien a, b positive ganze Zahlen mit $a \geq b$. Setze $a_0 = a$, $a_1 = b$ und erhalte rekursiv ganze Zahlen a_i, $i \geq 2$, mit*

$$a_{i-2} = q_{i-1}a_{i-1} + a_i, \quad q_{i-1}, a_i \in \mathbb{Z},\ 0 \leq a_i < a_{i-1}.$$

Dann gibt es einen ersten Index $k \geq 1$ mit $a_{k+1} = 0$, und es ist $ggT(a, b) = a_k$.

Beweis. Wegen $b = a_1 > a_2 > a_3 > \ldots \geq 0$ und $a_1 = b \geq 1$ gibt es einen ersten Index k mit $a_k > 0$ und $a_{k+1} = 0$. Dann ist $a_{k-1} = q_k a_k$, also ist a_k ein Teiler von a_{k-1}. Nun ist $a_{i-2} = q_{i-1}a_{i-1} + a_i$, daher ist a_k ein Teiler von a_{i-2}, falls a_k Teiler von a_{i-1} und a_i ist. Induktiv folgt damit $a_k|a_1$ und $a_k|a_0$. Also ist a_k ein gemeinsamer Teiler von a und b.

Sei umgekehrt c ein gemeinsamer Teiler von a und b, also von a_0 und a_1. Nun ist $a_i = a_{i-2} - q_{i-1}a_{i-1}$, also ist c ein Teiler von a_i, falls c Teiler von a_{i-2} und a_{i-1} ist. Induktiv folgt $c|a_k$. □

3.20. KOROLLAR. *Seien $a, b \in \mathbb{Z}$. Dann gibt es ganze Zahlen l, m mit*

$$\mathrm{ggT}(a, b) = la + mb.$$

Beweis. Ohne Beschränkung der Allgemeinheit dürfen wir $a \geq b > 0$ annehmen. Wir betrachten den Euklidischen Algorithmus aus Satz 3.19: Nach den Bezeichnungen dort ist

$$(*) \qquad \mathrm{ggT}(a, b) = a_k = a_{k-2} - q_{k-1}a_{k-1}.$$

Wegen $a_i = a_{i-2} - q_{i-1}a_{i-1}$ können wir auf der rechten Seite von $(*)$ rekursiv jeweils die Zahl a_i mit dem höchsten Index durch die entsprechende Kombination der zwei vorherigen Zahlen, nämlich a_{i-1} und a_{i-2} ersetzen. Auf diese Weise erhalten wir schliesslich $a_k = la + mb$ mit geeigneten ganzen Zahlen l und m. □

[3] Marin Mersenne (1588–1648).

4. Rechnen mit Kongruenzen

Sei m eine beliebige, aber fest gewählte ganze Zahl. Wir setzen

$$\mathbb{Z}m = \{km \mid k \in \mathbb{Z}\} = m\,\mathbb{Z}, \tag{4.1}$$

die Menge der ganzen Vielfachen von m. Für $a, b \in \mathbb{Z}$ gilt offenbar $a - b \in \mathbb{Z}m$ genau dann, wenn m ein Teiler von $a - b$ ist. Wir sagen dann, *a ist kongruent zu b modulo m* und schreiben $a = b$ modulo m oder kurz: $a = b \bmod m$. Wenn $m > 0$ ist, so ist offenbar $a = b \bmod m$ genau dann, wenn a und b bei der Division durch m denselben Rest im Sinne von Satz 3.2 haben.

4.2. Satz. *Die Relation $a = b \bmod m$ ist eine Äquivalenzrelation, siehe Definition 15.1. Mit anderen Worten, für alle $a, b, c \in \mathbb{Z}$ ist*

(1) $a = a \bmod m$,
(2) $a = b \bmod m \;\Rightarrow\; b = a \bmod m$,
(3) $a = b \bmod m$ *und* $b = c \bmod m \;\Rightarrow\; a = c \bmod m$. □

4.3. Rechenregeln. *Seien $a, a', b, b' \in \mathbb{Z}$ mit $a = a' \bmod m$ und $b = b' \bmod m$. Dann ist*

(1) $a + b = a' + b' \bmod m$.
(2) $ab = a'b' \bmod m$.

Beweis. Es gibt $k, l \in \mathbb{Z}$ mit $a - a' = km$ und $b - b' = lm$. Daher ist

$$(a + b) - (a' + b') = (a - a') + (b - b') = km + lm = (k + l)m,$$

damit (1). Ferner ist

$$ab - a'b' = (a - a')b + a'(b - b') = kmb + a'lm = (kb + a'l)m,$$

damit (2). □

Zu $a \in \mathbb{Z}$ setzen wir jetzt

$$[a] := a + \mathbb{Z}m = \{a + km | k \in \mathbb{Z}\} = \{b \mid a = b \bmod m\}, \tag{4.4}$$

die *Kongruenzklasse* bzw. *Restklasse* von $a \bmod m$ [4]. Es gilt immer $a \in [a]$, insbesondere ist $[a] \neq \emptyset$. Wichtig in allen Lebenslagen: Das Rechnen mit Kongruenzklassen ist einfacher als das Rechnen mit ganzen Zahlen.

4.5. Beispiele. 1) Die Uhrzeit: Das Zifferblatt der Uhr zeigt die Stunden von 1 bis 12. Wenn es jetzt 3 Uhr ist, so ist es in 17 Stunden 8 Uhr: In 12 Stunden dreht sich der Zeiger der Uhr einmal voll um seine Achse, er zeigt dann wieder auf die 3. Es bleiben 5 Stunden und $3 + 5 = 8$. Bei der Uhrzeit rechnen wir modulo 12. [5] Also ist

$$[3] = \{\ldots, -21, -9, 3, 15, 27, \ldots\} \text{ und } [17] = \{\ldots, -19, -7, 5, 17, 29, \ldots\}.$$

Es ist eine sehr empfehlenswerte Übung, die Definitionen und Resultate dieses Abschnitts im Beispiel der Uhrzeit durchzuspielen.

[4] Ich habe die Notation aus der Vorlesung geändert, $\overline{a}$ wird zu $[a]$.

[5] Wenn man noch den Unterschied von Vormittag und Nachmittag berücksichtigt, dann rechnet man modulo 24.

2) Die Winkelmessung in Grad: Der volle Kreis hat 360 Grad, bei der Winkelmessung rechnet man deshalb modulo 360.

3) Rechenkünstler berechnen auf Zuruf des Datums den zugehörigen Wochentag. Zur Trickkiste gehört dabei Rechnen modulo 7. Für Komplikationen sorgen die unterschiedlichen Längen der Monate, Schalttage und die Gregorianische Kalenderreform. Trotzdem ist das Verfahren nicht besonders schwierig, bei etwas Talent im Kopfrechnen kann man es in wenigen Stunden bis zur Gesellenprüfung bringen. Das Verfahren wird auf der Heimatseite von Gregor Weingart erklärt:

www.math.uni-bonn.de/people/gw.

4.6. LEMMA. *Seien $a, b \in \mathbb{Z}$. Dann ist $a = b \bmod m \Leftrightarrow [a] = [b] \Leftrightarrow [a] \cap [b] \neq \emptyset$.*

Beweis. Wir führen einen Ringschluss durch, wir wiederholen dabei das Argument aus dem allgemeineren Satz 15.2: Sei $a = b \bmod m$. Dann gibt es ein $l \in \mathbb{Z}$ mit $a - b = lm$. Für $k \in \mathbb{Z}$ ist dann $a + km = b + lm + km = b + (l + k)m$, mithin $[a] \subset [b]$. Wir können die Rollen von a und b vertauschen, also folgt ebenso $[b] \subset [a]$, damit insgesamt $[a] = [b]$. Daher folgt die zweite Eigenschaft aus der ersten.

Falls die zweite Eigenschaft gilt, so auch die dritte, den Kongruenzklassen sind niemals leer.

Sei nun $c \in [a] \cap [b]$. Dann gibt es $k, l \in \mathbb{Z}$ mit $c = a + km = b + lm$. Also ist $a - b = (l - k)m$ und damit $a = b \bmod m$. □

Die Menge der Kongruenzklassen nennen wir $\mathbb{Z}$ *modulo* m und schreiben sie

$$\mathbb{Z}/m = \mathbb{Z}/m\mathbb{Z} := \{[a] \mid a \in \mathbb{Z}\}. \tag{4.7}$$

Falls $x \in \mathbb{Z}/m$ und a eine ganze Zahl ist mit $a \in x$, so ist $x = [a]$. Dann heisst a *Repräsentant* von x. Für $m > 0$ gilt offenbar: Jede der Klassen in $\mathbb{Z}/m$ ist durch genau eine der Zahlen $0, \ldots, m - 1$ repräsentiert, $\mathbb{Z}/m$ hat damit m Elemente, $|\mathbb{Z}/m| = m$.

Wir definieren jetzt Addition und Multiplikation von Kongruenzklassen: Seien $x, y \in \mathbb{Z}/m$. Wähle $a, b \in \mathbb{Z}$ mit $a \in x$, $b \in y$. Setze

$$x + y := [a + b], \quad xy := [ab]. \tag{4.8}$$

Die rechten Seiten sind wohldefiniert, d.h., unabhängig von der Wahl der Repräsentanten. Falls nämlich $a' \in x$ und $b' \in y$ weitere Repräsentanten sind, so ist $a = a' \bmod m$ und $b = b' \bmod m$, also $a+b = a+b' \bmod m$ und $ab = a'b' \bmod m$ mit den Rechenregeln 4.3. Mithin $[a + b] = [a' + b']$ und $[ab] = [a'b']$ nach Lemma 4.6.

Das Rezept für Additon und Multiplikation lautet also: Wähle Repräsentanten, bilde ihre Summe bzw. ihr Produkt, nehme die Kongruenzklasse des Resultats. Damit ist nicht erstaunlich, dass die Rechenregeln für das Rechnen mit ganzen Zahlen weiterhin Bestand haben.

4.9. LEMMA. *Für alle $x, y, z \in \mathbb{Z}/m$ gilt:*

(1) $(x + y) + z = x + (y + z)$ *und* $(xy)z = x(yz)$.
(2) $x + y = y + x$ *und* $xy = yx$
(3) $x + [0] = [0] + x = x$ *und* $x \cdot [1] = [1] \cdot x = x$.
(4) *Falls $a \in x$, so ist* $[-a] + x = x + [-a] = [0]$.

Beweis. Wir nützen aus, dass die entsprechenden Regeln für das Rechnen mit ganzen Zahlen gelten. Seien $a \in x, b \in y, c \in z$ Repräsentanten. Dann ist

$$\begin{aligned}(x+y)+z &= ([a]+[b])+[c] = [a+b]+[c] \\ &= [(a+b)+c] = [a+(b+c)] = x+(y+z).\end{aligned}$$

Den Beweis der restlichen Regeln lasse ich als Übung. □

Ich komme jetzt zu der anfangs angekündigten Anwendung in der Kryptographie, zum *RSA-Verfahren*. Das Verfahren beruht auf einem Satz der Mathematiker Fermat[6] und Euler[7].

4.10. DEFINITION. Wir nennen $a, m \in Z$ *teilerfremd*, falls $\mathrm{ggT}(a,m) = 1$ ist. Für $m \geq 1$ sei $\varphi(m)$ die Anzahl der zu m teilerfremden Zahlen aus $\{1,\dots,m\}$. Wir nennen φ die *Eulersche φ-Funktion*.

Zum Beispiel ist $\varphi(1) = 1, \varphi(2) = 1, \varphi(3) = 2, \varphi(4) = 2, \varphi(5) = 4, \varphi(6) = 2, \dots$

4.11. SATZ (Euler-Fermat). *Seien a, m teilerfremde positive ganze Zahlen. Dann ist*

$$a^{\varphi(m)} = 1 \bmod m.$$

Dieser Satz enthält den *Kleinen Satz von Fermat* als Spezialfall: Fermat bewies Satz 4.11 unter der Voraussetzung, dass m prim ist. Dann ist $\varphi(m) = m - 1$.

In [RU], Abschnitt 1 in Kapitel 5, findet man einen zahlentheoretischen Beweis dieses Satzes. Wir werden in unserem Beweis Hilfsmittel aus der Gruppentheorie verwenden und verschieben den Beweis daher an das Ende von Abschnitt 6.

Nun zum mathematischen Aspekt des RSA-Verfahrens: Man verschlüssele Nachrichten zunächst auf eine geeignete Weise mit natürlichen Zahlen. Solche Nachrichten werden in der Praxis eine gewisse Länge nicht überschreiten.

Das RSA-Verfahren ist asymmetrisch, das heisst, Sender und Emfänger sind nicht gleichberechtigt. Der Emfänger wählt Primzahlen p und q, die sehr groß sind und auf jeden Fall (viel) größer als die Nachrichten. Er setzt $m = pq$. Dann ist

$$\varphi(m) = (p-1)(q-1) = m + 1 - (p+q),$$

wie man durch Abzählen leicht feststellt. Der Empfänger wählt ausserdem eine weitere Zahl e, die teilerfremd zu $\varphi(m)$ ist. Der Name e für diese Zahl steht für *encryption*. Der *public key* des Emfängers besteht aus den Zahlen m und e. Diese veröffentlicht er, die Sender der Nachrichten benützen sie, um ihre Nachrichten zu verschlüsseln.

Der *private key* des Empfängers besteht aus einer ganzen Zahl d mit $ed = 1 \bmod \varphi(m)$. Es gibt also eine ganze Zahl k mit $ed = k\varphi(m) + 1$.

Der Sender berechnet den Rest c von a^e bei der Division durch m im Sinne von Satz 3.2. Dann ist $0 \leq c < m$. Diesen Rest sendet er an den Empfänger. Dieser berechnet dann den Rest von c^d bei der Division durch m. Sender und Empfänger

[6] Pierre de Fermat (1601–1665).

[7] Leonhard Euler (1707–1783).

verwenden für ihre Berechnungen *schnelle* Algorithmen wie etwa die in Aufgabe 4 von Übungsblatt 4 vorgestellten Verfahren. In $\mathbb{Z}/m$ gilt

$$[c^d] = [a^{ed}] = [a^{k\varphi(m)+1}] = [(a^{\varphi(m)})^k] \cdot [a].$$

Nun ist a kleiner als p und q und damit teilerfremd zu m. Mit dem Satz 4.11 von Euler-Fermat folgt deshalb $a^{\varphi(m)} = 1 \bmod m$. Also ist

$$[(a^{\varphi(m)})^k] = [1] \quad \text{und damit} \quad [(a^{\varphi(m)})^k] \cdot [a] = [a].$$

Daher ist $a = c^d \bmod m$. Nun ist $0 \leq a < m$, also ist a der Rest von c^d bei der Division durch m. Der Emfänger hat die Nachricht a damit entschlüsselt.

Der Punkt des Verfahrens ist, dass es Dritten mit den bisher bekannten Verfahren nicht möglich ist, p und q aus m in kurzer Zeit zurückzugewinnen, jedenfalls nicht, wenn p und q groß genug sind. Damit bleibt Dritten auch $\varphi(m)$ unbekannt.

5. GRUPPEN

5.1. DEFINITION. Eine *Gruppe* ist eine Menge G zusammen mit einer Verknüpfung [8]

$$G \times G \to G, \quad (g,h) \mapsto g * h,$$

so dass die folgenden drei Eigenschaften erfüllt sind:

(1) Für alle $g, h, k \in G$ ist $(g * h) * k = g * (h * k)$ (Assoziativgesetz).
(2) Es gibt ein Element $e \in G$ mit $g * e = e * g = g$ für alle $g \in G$. Man nennt e das *neutrale Element* von G.
(3) Für jedes $g \in G$ gibt es ein Element $h \in G$ mit $g * h = h * g = e$. Man nennt h das zu g *inverse Element*.

Falls weiter gilt $g * h = h * g$ für alle $g, h \in G$, so nennt man die Gruppe *kommutativ* oder *Abelsch* [9].

5.2. BEISPIELE. 0) Die einer Gruppe zugrunde liegende Menge ist nicht leer, die Gruppe enthält ja das neutrale Element. Auf einer Menge mit einem Element gibt es nur eine mögliche Verknüpfung, diese erfüllt die Gruppenaxiome. Damit haben wir das nullte Beispiel: die *triviale Gruppe* $G = \{e\}$.

1) Die Menge $\mathbb{Z}$ mit der Addition als Verknüpfung, $(a,b) \mapsto a + b$. Neutrales Element ist die 0, das zu a inverse Element ist $-a$. Diese Gruppe ist kommutativ. $(\mathbb{Q}, +), (\mathbb{R}, +), (\mathbb{C}, +)$ sind ebenfalls kommutative Gruppen, $(\mathbb{N}, +)$ ist keine Gruppe.

2) Sei $m \in \mathbb{Z}$ und die Additon in $\mathbb{Z}/m$ wie in (4.8) definiert. Nach Lemma 4.9 ist dann $(\mathbb{Z}/m, +)$ eine kommutative Gruppe.

3) Sei X eine nicht leere Menge. Eine *Permutation* von X ist eine bijektive Abbildung $X \to X$. Die *symmetrische Gruppe* S_X besteht aus allen Permutationen von X zusammen mit der Komposition als Verknüpfung. Neutrales Element von S_X ist die identische Abbildung von X, inverses Element zu $f \in S_X$ ist die Umkehrabbildung f^{-1} von f. Für $|X| \geq 3$ ist S_X nicht Abelsch. (Siehe Übungsblatt 5, Aufgabe 3.)

5.3. KONVENTION. Man schreibt die Verknüpfung in einer Gruppe G zumeist als Multiplikation, also die Verknüpfung von $g, h \in G$ als $g \cdot h$ bzw. einfach als gh und nennt das Resultat das *Produkt* von g und h. Es ist dann manchmal zweckmässig, das neutrale Element mit 1 und das inverse Element von g mit g^{-1} zu bezeichnen.

Bei Abelschen Gruppen schreibt man die Verknüpfung allerdings häufig als Addition, also als $g + h$, und nennt dies die *Summe* von g und h. Jetzt ist es manchmal zweckmässig, das neutrale Element mit 0 und das inverse Element zu g mit $-g$ zu bezeichnen. In einer additiven Gruppe schreiben wir auch $x - y$ statt $x + (-y)$.

In der Definiton wurde das Symbol $*$ gewählt, um nicht gleich zu Anfang mit den verschiedenen Symbolen für neutrales und inverses Element in den Beispielen Verwirrung zu stiften. Wir werden ab jetzt die multiplikative Schreibweise benützen, für Abelsche Gruppen auch die additive.

5.4. SATZ (Kürzungsregel). *Sei G eine Gruppe. Seien $g, h, k \in G$, und sei $gk = hk$ oder $kg = kh$. Dann ist $g = h$.*

[8] *Verknüpfung* ist hier nur ein anderer Name für Abbildung.

[9] Niels Henrik Abel (1802–1829).

Beweis. Es gelte $gk = hk$. Multiplikation mit k^{-1} von rechts liefert

$$g = ge = g(kk^{-1}) = (gk)k^{-1} = (hk)k^{-1} = h(kk^{-1}) = he = h,$$

also $g = h$. Analog argumentiert man im zweiten Fall. □

5.5. KOROLLAR. *(1) Falls $g \in G$ ist und $gg = g$, so ist $g = e$.*
(2) Falls $g, h \in G$ sind und $gh = e$ ist, so ist $g = h^{-1}$.

Beweis. In (1), ersetze die rechte Seite durch ge, in (2) durch $h^{-1}h$. □

5.6. DEFINITION. Sei G eine Gruppe. Wir nennen $H \subset G$ eine *Untergruppe*, wenn

(1) $e \in H$ ist und
(2) mit g und h auch gh und g^{-1} in H sind.

Anders gesagt: Mit der gegebenen Verknüpfung wird eine Untergruppe selber zur Gruppe.

5.7. BEISPIELE. 0) Die *triviale Untergruppe* $H = \{e\}$ und $H = G$ sind Untergruppen.

1) $\mathbb{Z}m \subset \mathbb{Z}$, $\mathbb{Z} \subset \mathbb{Q}$ und $\mathbb{Q} \subset \mathbb{R}$ sind Untergruppen der additiven Gruppen $\mathbb{Z}$ bzw. $\mathbb{Q}$ bzw. $\mathbb{R}$.

3) In der additiven Gruppe $\mathbb{Z}/6$ ist $H = \{[0], [2], [4]\}$ eine Untergruppe.

Sei G eine Gruppe und H eine Untergruppe. Wir schreiben $g = h$ modulo H, oder kurz: $g = h \bmod H$, wenn $g^{-1}h \in H$ ist.

5.8. LEMMA. *Die Relation $g = h \bmod H$ ist eine Äquivalenzrelation.*

Beweis. Für alle $g \in G$ gilt $g = g \bmod H$ denn $g^{-1}g = e \in H$. Mit $g = h \bmod H$ ist auch $h = g \bmod H$, denn mit $g^{-1}h \in H$ ist auch $h^{-1}g = (g^{-1}h)^{-1} \in H$. Mit $g = h \bmod H$ und $h = k \bmod H$ ist auch $g = k \bmod H$, denn mit $g^{-1}h \in H$ und $h^{-1}k \in H$ ist auch $g^{-1}k = (g^{-1}h)(h^{-1}k) \in H$. □

Die Äquivalenzklasse $[g]$ von g nennen wir *Restklasse* oder *Kongruenzklasse* von $g \bmod H$, sie ist

$$[g] = gH = \{gh \mid h \in H\}. \tag{5.9}$$

Insbesondere besteht die Restklasse des neutralen Elementes genau aus der Untergruppe H. Nach der Kürzungsregel 5.4 für Gruppen ist die Abbildung $H \to gH$, $h \mapsto gh$, bijektiv für alle $g \in G$. Das folgende Lemma folgt aus Satz 15.2; als Übung versuche der Leser, das Argument aus dem Beweis des analogen Lemmas 4.6 zu übertragen.

5.10. LEMMA. *Für $g, h \in G$ ist $g = h \bmod H \Leftrightarrow [g] = [h] \Leftrightarrow [g] \cap [h] \neq \emptyset$.* □

5.11. DEFINITION. Sei G eine Gruppe. Dann nennen wir $\operatorname{ord} G := |G|$ die *Ordnung von G*.

5.12. SATZ (Lagrange). [10] *Sei G eine Gruppe endlicher Ordnung und $H \subset G$ eine Untergruppe. Dann ist $\operatorname{ord} H$ ein Teiler von $\operatorname{ord} G$.*

Beweis. Die Restklassen von $G \bmod H$ haben genau $|H| = \operatorname{ord} H$ Elemente, ihre Vereinigung ist G, und sie sind paarweise disjunkt. □

[10] Joseph Louis Lagrange (1736–1813).

Sei G eine (multiplikativ geschriebene) Gruppe und $g \in G$. Wir definieren die Potenzen von g wie folgt: $g^0 := e$, $g^1 := g$ und, für $n \geq 2$,

$$g^n := gg^{n-1} \text{ bzw. } g^{-n} := g^{-1}g^{-(n-1)}. \tag{5.13}$$

Also $g^1 = g$, $g^2 = gg$, $g^3 = g(gg)$, $g^{-2} = g^{-1}g^{-1}, \ldots$ Dann ist

$$g^{k+l} = g^k g^l \quad \text{für alle } k, l \in \mathbb{Z}. \tag{5.14}$$

In einer additiven Gruppe ist es nicht angebracht, von den Potenzen eines Elementes zu sprechen. Die analoge Definition führt hier zu den *Vielfachen*

$$0 \cdot a := 0,\ n \cdot a := a + (n-1) \cdot a,\ (-n) \cdot a = -a + (1-n) \cdot a$$

für alle $n \geq 1$.

Zurück zur multiplikativ geschriebenen Gruppe G und dem gegebenen Element $g \in G$: Aus (5.14) folgt, dass $\langle g \rangle := \{g^k | k \in \mathbb{Z}\}$ eine Untergruppe von G ist, *die von g erzeugte Untergruppe*. Wir nennen $\operatorname{ord} g := |\langle g \rangle|$ *die Ordnung von g*.

Es gibt zwei Möglichkeiten: Entweder sind die Potenzen g^k, $k \in \mathbb{Z}$, alle paarweise verschieden, und dann ist $\operatorname{ord} g = \infty$. Oder es gibt $k < l$ mit $g^k = g^l$, also $g^{l-k} = e$ mit $l - k > 0$. Dann gibt es ein kleinstes $n \geq 1$ mit $g^n = e$. Für dieses sind die Potenzen $e, g, \ldots, g^{n-1}$ von g paarweise verschieden. Für $m \in \mathbb{Z}$ beliebig ist nun $m = qn + r$ mit $0 \leq r < n$, also $g^m = g^{qn+r} = (g^n)^q g^r = g^r$, also ist dann $\langle g \rangle = \{e, g, \ldots, g^{n-1}\}$ und $\operatorname{ord} g = n$. Mit anderen Worten: Falls $\operatorname{ord} g$ endlich ist, so ist $\operatorname{ord} g$ das kleinste $n \geq 1$ mit $g^n = e$.

5.15. KOROLLAR. *Sei G eine Gruppe endlicher Ordnung und $g \in G$. Dann ist* $\operatorname{ord} g$ *ein Teiler von* $\operatorname{ord} G$. *Insbesondere gilt $g^{\operatorname{ord} G} = e$ für alle $g \in G$.*

Dies ist das gruppentheoretische Resultat, welches wir im Beweis des Satzes 4.11 von Euler-Fermat benützen werden.

6. RINGE

6.1. DEFINITION. Ein *Ring* ist eine Menge R zusammen mit zwei Verknüpfungen,

$$R \times R \to R, \quad (x,y) \mapsto x+y \qquad \textit{(Addition)},$$
$$R \times R \to R, \quad (x,y) \mapsto x \cdot y \qquad \textit{(Multiplikation)},$$

so dass die folgenden drei Eigenschaften erfüllt sind:

(1) $(R,+)$ ist eine kommutative Gruppe.
(2) Für alle $x,y,z \in R$ ist $(xy)z = x(yz)$. (Assoziativgesetz)
(3) Für alle $x,y,z \in R$ ist $(x+y)z = xz+yz$ und $x(y+z) = xy+xz$. (Distributivgesetze)

Falls noch zusätzlich $xy = yx$ für alle $x,y \in R$ ist, so heißt der Ring kommutativ.

Das neutrale Element der Addition in einem Ring R bezeichnen wir meist mit 0. Für $x,y \in R$ schreiben wir auch xy statt $x \cdot y$. Für alle $x \in R$ gilt $0 \cdot x = x \cdot 0 = 0$, denn $0 \cdot x = (0+0) \cdot x = 0 \cdot x + 0 \cdot x$.

6.2. DEFINITION. Sei R ein Ring. Dann heisst $S \subset R$ *Unterring* von R, wenn

(1) $(S,+)$ Untergruppe von $(R,+)$ ist und
(2) mit x,y auch xy in S ist.

Eine Unterring zusammen mit den eingeschränkten Verknüpfungen ist damit selber ein Ring.

6.3. BEISPIELE. 0) Der *triviale Ring* $R = \{0\}$. Jeder Ring enthält den trivialen Ring als Unterring.

1) Zusammen mit der üblichen Addition und Multiplikation sind $\mathbb{Z}$, $\mathbb{Q}$ und $\mathbb{R}$ kommutative Ringe. $\mathbb{Z}$ ist Unterring von $\mathbb{Q}$ und $\mathbb{R}$, $\mathbb{Q}$ ist Unterring von $\mathbb{R}$.

2) Nach Lemma 4.9 ist $\mathbb{Z}/m$ zusammen mit der in (4.8) definierten Addition und Multiplikation ein kommutativer Ring.

Ein Ring R heißt *Ring mit Eins* , falls die Multiplikation ein neutrales Element hat. Dieses bezeichnen wir zumeist mit 1; für alle $x \in R$ gilt also $1 \cdot x = x \cdot 1 = x$.

6.4. LEMMA. *Sei R ein Ring mit Eins und $R \neq \{0\}$. Dann ist $1 \neq 0$.*

Beweis. Sei $1 = 0$. Sei $x \in R$. Dann ist $x = 1 \cdot x = 0 \cdot x = 0$, also $x = 0$ und damit $R = \{0\}$. □

6.5. BEISPIELE. 1) $\mathbb{Z}$, $\mathbb{Q}$ und $\mathbb{R}$ sind Ringe mit Eins.

2) $\mathbb{Z}/m$ ist Ring mit Eins, das neutrale Element der Multiplikation ist $[1]$.

Sei R ein Ring mit Eins. Dann heißt $x \in R$ *Einheit* in R, wenn es $x', x'' \in R$ gibt, so dass $x'x = xx'' = 1$ ist. Wir nennen

$$R^* := \{x \in R \mid x \text{ ist Einheit in } R\} \tag{6.6}$$

die *Gruppe der Einheiten* von R. Es ist $1 \in R^*$, insbesondere $R^* \neq \emptyset$. Mit $x'x = xx'' = 1$ folgt $x' = x'(xx'') = (x'x)x'' = x''$, also $x' = x''$ in der Definition von R^*. Wir nennen x' das zu $x \in R^*$ *(multiplikativ) inverse Element* und bezeichnen es mit x^{-1}. Mit $x,y \in R^*$ ist auch $xy \in R^*$, denn $(xy)(y^{-1}x^{-1}) = x(yy^{-1})x^{-1} = xx^{-1} = 1$

und analog $(y^{-1}x^{-1})(xy) = 1$. Wir sehen damit auch, dass $y^{-1}x^{-1}$ das Inverse zu xy ist. Als Konsequenz der Diskusion folgt der

6.7. SATZ. *$(R^*, \cdot)$ ist Gruppe mit neutralem Element 1.*

6.8. BEISPIELE. $\mathbb{Z}^* = \{1, -1\}$, $\mathbb{Q}^* = \mathbb{Q} \setminus \{0\}$, $\mathbb{R}^* = \mathbb{R} \setminus \{0\}$.

6.9. SATZ. *Sei $m \in \mathbb{Z}$. Dann ist $(\mathbb{Z}/m)^* = \{[a] \mid \mathrm{ggT}(a, m) = 1\}$. Neutrales Element ist $[1]$.*

Beweis. Falls $\mathrm{ggT}(a, m) = 1$ ist, so gibt es $k, l \in \mathbb{Z}$ mit $ka + lm = 1$. Damit folgt $[ka] = [1]$, also $[k] \cdot [a] = [1]$. Daher ist $[k]$ ein multiplikatives Inverses von $[a]$. Falls umgekehrt $[k] \cdot [a] = [1]$ ist, so gibt es ein $l \in \mathbb{Z}$ mit $ka + lm = 1$. Dann ist aber $\mathrm{ggT}(a, m) = 1$. □

6.10. BEISPIELE. 1) Für $m = 0$ ist $[a] = \{a\}$ für alle $a \in \mathbb{Z}$ und $(\mathbb{Z}/m)^* = \{\pm[1]\}$.

2) Für $m = 1$ ist $[a] = \mathbb{Z}$ für alle $a \in \mathbb{Z}$, also $[1] = [0]$, $\mathbb{Z}/m = \{[0]\}$ und $(\mathbb{Z}/m)^* = \{[0]\}$.

3) Für $m = 4$ gilt: $(\mathbb{Z}/4)^* = \{[1], [3]\}$.

4) Für $m > 1$ prim ist $(\mathbb{Z}/m)^* = \{[1], \ldots, [m-1]\}$, jedes Element $x \neq [0]$ in $\mathbb{Z}/m$ hat daher ein multiplikatives Inverses.

5) Sei $m = pq$ mit p, q prim und $p \neq q$. Dann ist

$$(\mathbb{Z}/m)^* = \{[0], \ldots, [m-1]\} \setminus \{[0], [p], \ldots, [(q-1)p], [q], \ldots, [(p-1)q]\}.$$

Wir erinnern uns an den Satz 4.11 von Euler und Fermat: Seien a, m teilerfremde positive ganze Zahlen. Dann ist

$$a^{\varphi(m)} = 1 \bmod m.$$

Jetzt stehen alle Hilfsmittel für den Beweis bereit.

Beweis von Satz 4.11. Weil a und m teilerfremd sind, ist $[a] \in (\mathbb{Z}/m)^*$. Nach Definiton von φ und Satz 6.9 ist $\mathrm{ord}(\mathbb{Z}/m)^* = \varphi(m)$. Mit Korollar 5.15 folgt $[a]^{\varphi(m)} = [1]$, das ist aber gerade die Behauptung. □

7. KÖRPER, KOMPLEXE ZAHLEN

7.1. DEFINITION. Sei $(R, +, \cdot)$ ein kommutativer Ring mit Eins. Falls dann $1 \neq 0$ und die Gruppe der Einheiten $R^* = R \setminus \{0\}$ ist, so nennt man $(R, +, \cdot)$ auch einen *Körper*.
7.2. BEMERKUNG. Verzichtet man auf die Kommutativität von R, so spricht man auch von einem *Divisionsring*. Beispiele für Divisionsringe werden sie später kennenlernen.
7.3. DEFINITION. Sei K ein Körper. Dann heisst $L \subset K$ *Unterkörper* von K, wenn

(1) $(L, +)$ eine Untergruppe von $(K, +)$ und
(2) $(L \setminus \{0\}, \cdot)$ eine Untergruppe von $(K \setminus \{0\}, \cdot)$ ist.

Dann heisst K auch *Erweiterungskörper* von L.
7.4. BEISPIELE. 1) Mit der üblichen Addition und Multiplikation sind $\mathbb{Q}$ und $\mathbb{R}$ Körper, $\mathbb{Z}$ aber nicht; $\mathbb{Q}$ ist Unterkörper von $\mathbb{R}$.

2) Aus Satz 6.9 folgt, dass $\mathbb{Z}/m$ mit der in (4.8) definierten Addition und Multiplikation genau dann ein Körper ist, wenn $|m|$ eine Primzahl ist.

3) Sei $m > 0$ kein Quadrat in $\mathbb{Q}$, zum Beispiel $m = 2$. Sei $\sqrt{m} \in \mathbb{R}$ eine (beliebige) der beiden reellen Quadratwurzeln aus m und $\mathbb{Q}[\sqrt{m}] = \{a + b\sqrt{m} \mid a, b \in \mathbb{Q}\}$. Dann ist $\mathbb{Q}[\sqrt{m}]$ Unterkörper von $\mathbb{R}$ und Erweiterungskörper von $\mathbb{Q}$. (Vergleiche Übungsblatt 7, Aufgabe 6.)

Neben den Körpern $\mathbb{Q}$ der rationalen und $\mathbb{R}$ der reellen Zahlen ist der *Körper der komplexen Zahlen* der dritte im Dreigestirn der klassischen Körper. Die zugrunde liegende Menge ist $\mathbb{C} := \mathbb{R}^2 = \{(x, y) \mid x, y \in \mathbb{R}\}$, Addition und Multiplikation sind definiert durch

$$(x, y) + (x', y') := (x + x', y + y') \quad \text{und} \quad (x, y) \cdot (x', y') := (xx' - yy', xy' + yx').$$

Sei $(x, y) \in \mathbb{C}$, $(x, y) \neq (0, 0)$. Dann ist $x^2 + y^2 > 0$ und man rechnet leicht nach, dass

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

das zu (x, y) muliplikativ inverse Element ist. Den Nachweis der übrigen Körperaxiome überlassen wir als Übung.
7.5. KONVENTION. $\mathbb{R} \times \{0\} \subset \mathbb{C}$ soll mit $\mathbb{R}$ gleichgesetzt werden, d.h., wir schreiben x statt $(x, 0)$. Auf diese Weise fassen wir $\mathbb{R}$ als Unterkörper von $\mathbb{C}$ auf. Statt $(0, 1)$ schreiben wir i, so dass $(0, y) = (0, 1) \cdot (y, 0) = iy = yi$ und $(x, y) = x + iy = x + yi$.
7.6. DEFINITION. Sei $z = x + iy \in \mathbb{C}$ mit $x, y \in \mathbb{R}$. Dann nennen wir

(1) $\bar{z} = x - iy$ die zu z *konjugierte* komplexe Zahl,
(2) $x = \frac{1}{2}(z + \bar{z})$ den *Realteil* von z,
(3) $y = \frac{1}{2i}(z - \bar{z})$ den *Imaginärteil* von z und
(4) $|z| = \sqrt{x^2 + y^2} = \sqrt{z\bar{z}}$ den *Betrag* oder die *Länge* von z.

Hierbei ist natürlich die nichtnegative Wurzel aus der nichtnegativen Zahl unter dem Wurzelzeichen gemeint.

7.7. RECHENREGELN. *Für alle $w, z \in \mathbb{C}$ gilt*

(1) $\overline{\bar{z}} = z$
(2) $\overline{w+z} = \bar{w} + \bar{z}$, $\overline{wz} = \bar{w}\bar{z}$;
(3) *Für $z \neq 0$ ist $z^{-1} = \bar{z}/z\bar{z}$; mithin* $(\bar{z})^{-1} = \overline{(z^{-1})}$;
(4) $z = \operatorname{Re} z \iff \operatorname{Im} z = 0 \iff z = \bar{z} \iff z \in \mathbb{R}$,
$z = i \operatorname{Im} z \iff \operatorname{Re} z = 0 \iff z = -\bar{z} \iff z \in i\mathbb{R}$;
(5) $|z| > 0$ *falls* $z \neq 0$;
(6) $|wz| = |w| \cdot |z|$;
(7) $|w+z| \leq |w| + |z|$, *die* Dreiecksungleichung;
(8) $|\operatorname{Re} z| \leq |z|$, $|\operatorname{Im} z| \leq |z|$, $|\bar{z}| = |z|$.

Beweis. Wir zeigen die Behauptungen (6) und (7):

$$|wz|^2 = (wz)(\bar{w}\bar{z}) = (w\bar{w})(z\bar{z}) = |w|^2|z|^2,$$

also (6). Ferner gilt

$$\begin{aligned} |w+z|^2 &= (w+z)(\bar{w}+\bar{z}) = w\bar{w} + w\bar{z} + z\bar{w} + z\bar{z} \\ &= |w|^2 + w\bar{z} + \overline{w\bar{z}} + |z|^2 = |w|^2 + 2\operatorname{Re}(w\bar{z}) + |z|^2 \\ &\leq |w|^2 + 2|w\bar{z}| + |z|^2 = (|w| + |z|)^2. \end{aligned}$$

Die restlichen Behauptungen bleiben als Übung. □

Wenn man die Rechenregeln der komplexen Zahlen überprüft, stellt man fest, dass sie den Regeln im Körper $\mathbb{Q}[\sqrt{m}]$ in Beipiel 3 aus 7.4 sehr ähnlich sind. Im Falle der komplexen Zahlen ist ja auch $i^2 = -1$, eine Zahl, die keine rationale oder reelle Wurzel hat. Nach dem gleichen Rezept betrachtet man in der Zahlentheorie die Ringe $\mathbb{Z}[\sqrt{m}]$ und Körper $\mathbb{Q}[\sqrt{m}]$, wobei $m \in \mathbb{Z}$ in $\mathbb{Q}$ kein Quadrat ist. Der Fall $m > 0$ wurde schon in Beipiel 3 aus 7.4 betrachtet. Für $m = -1$ erhält man den *Ring* bzw. den *Körper* der Gaußschen Zahlen [11],

$$\mathbb{Z}[i] = \{z \in \mathbb{C} \mid \operatorname{Re} z, \operatorname{Im} z \in \mathbb{Z}\}, \quad \mathbb{Q}[i] = \{z \in \mathbb{C} \mid \operatorname{Re} z, \operatorname{Im} z \in \mathbb{Q}\}.$$

Die Gruppe der Einheiten von $\mathbb{Z}[i]$ ist $(\mathbb{Z}[i])^* = \{\pm 1, \pm i\}$. Man vergleiche hierzu die Abschnitte 3.0.3 und 3.1. in [RU].

In der Vorlesung zur Analysis wird gezeigt, dass jedes Polynom *ungeraden Grades* mit reellen Koeffizienten eine relle Nullstelle hat. Dagegen gibt es keine reellen Quadratwurzeln aus negativen reellen Zahlen, der Körper $\mathbb{C}$ der komplexen Zahlen entsteht durch "Hinzufügen"dieser Quadratwurzeln. In den Übungen (Aufgabe 5, Übungsblatt 7) haben wir gesehen, dass jedes quadratische Polynom mit komplexen Koeffizienten Nullstellen in $\mathbb{C}$ hat. Tatsächlich hat aber jedes Polynom mit komplexen Koeffizienten komplexe Nullstellen:

[11] Carl Friedrich Gauß (1777–1855).

7.8. SATZ (Fundamentalsatz der Algebra). *Jede Gleichung*

$$a_n z^n + a_{n-1} z^{n-1} + \ldots + a_1 z + a_0 = 0$$

mit $n > 0$*, Koeffizienten* $a_0, \ldots, a_n \in \mathbb{C}$ *und* $a_n \neq 0$ *hat eine Lösung in* $\mathbb{C}$*.*

Erste richtige Beweise dieses Satzes stammen von Laplace [12] (1795) und Gauß (1799). Inzwischen gibt es sehr viele verschiedene Beweise, einige davon lernt man zum Beispiel in den Vorlesungen zur Funktionentheorie und Algebra kennen.

Ein weiteres Resultat, das wir hier nicht beweisen werden, stammt von Euler:

7.9. SATZ (Eulersche Formel). *Für alle* $z \in \mathbb{C}$ *ist* $e^{iz} = \cos z + i \sin z$*.*

Diese Formel ist sehr wichtig und sehr praktisch. Zum Beispiel folgen aus der Produktregel $e^{w+z} = e^w e^z$ die Additionstheoreme für Sinus und Kosinus direkt und ohne Mühe.

[12] Pierre Simon Laplace (1749–1827).

8. VEKTORRÄUME

Mit den Vektorräumen betritt nun schliesslich (und endlich) einer der Hauptakteure der linearen Algebra die Bühne.

Sei K ein Körper und 1 das neutrale Element der Multiplikation in K.

8.1. DEFINITION. Ein *Vektorraum* über K ist eine Menge V zusammen mit zwei Verknüpfungen,

$$V \times V \to V, \quad (v, w) \mapsto v + w \qquad \textit{(Addition)},$$
$$K \times V \to V, \quad (\alpha, v) \mapsto \alpha \cdot v \qquad \textit{(skalare Multiplikation)},$$

so dass die folgenden fünf Eigenschaften erfüllt sind:

(1) $(V, +)$ ist eine Abelsche Gruppe.
(2) Für alle $v \in V$ ist $1 \cdot v = v$.
(3) Für alle $\alpha, \beta \in K$ und $v \in V$ ist $\alpha \cdot (\beta \cdot v) = (\alpha \cdot \beta) \cdot v$.
(4) Für alle $\alpha, \beta \in K$ und $v \in V$ ist $(\alpha + \beta) \cdot v = \alpha \cdot v + \beta \cdot v$.
(5) Für alle $\alpha \in K$ und $v, w \in V$ ist $\alpha \cdot (v + w) = \alpha \cdot v + \alpha \cdot w$.

Die Elemente aus V nennen wir dann *Vektoren*, Elemente aus K *Skalare*. Das neutrale Element der Addition in V bezeichnen wir meist mit 0 und nennen es den *Nullvektor*.

Einige Bemerkungen zur Schreibweise sind angebracht. Sowohl der Körper K als auch die Abelsche Gruppe haben ein neutrales Element der Addition, beide bezeichnen wir in der Regel mit der Ziffer 0. Aus dem Zusammenhang muss man dann erschliessen, um welches der beiden neutralen Elemente es sich handelt. Das ist praktisch immer offensichtlich. Analoges gilt für die Addition in K und V bzw. die Multiplikation in K und skalare Multiplikation, die schon in der Definition oben mit den gleichen Symbolen bezeichnet werden.

Wir vereinbaren noch die üblichen Konventionen: Zum Beispiel verzichten wir meist auf den Punkt bei der skalaren Multiplikation. Wir schreiben auch meist $v - w$ statt $v + (-w)$.

8.2. BEISPIELE. (0) Der *triviale Vektorraum*: $V = \{0\}$ mit den einzig möglichen Verknüpfungen.

(1) Das *Standard-Beispiel*: Sei n eine ganze Zahl ≥ 1 und

$$V = K^n := \{(x_1, \ldots, x_n) \mid x_1, \ldots, x_n \in K\}$$

die Menge der n-Tupel von Elementen aus K. Zu definieren sind Addition und skalare Multiplikation. Seien dazu $x = (x_1, \ldots, x_n)$ und $y = (y_1, \ldots, y_n)$ in K^n und $\alpha \in K$. Wir setzen

$$x + y := (x_1 + y_1, \ldots, x_n + y_n) \quad \text{und} \quad \alpha x := (\alpha x_1, \ldots, \alpha x_n),$$

wobei auf den rechten Seiten in jeder Komponente jeweils die gegebene Addition und Multiplikation aus K benützt wird. Der Nullvektor ist das Tupel $0 = (0, \ldots, 0) \in K^n$, dessen Einträge alle $0 \in K$ sind.

Hilfreich und wichtig ist die geometrische Interpretation der beiden Verknüpfungen. In der reellen Ebene $\mathbb{R}^2$ ist sie wie folgt: Wir stellen uns den Vektor $x = (x_1, x_2)$

als Strecke mit Anfang im Nullvektor und Endpunkt (oder Spitze) in (x_1, x_2) vor. Dann entspricht αx der um α reskalierten Strecke mit Anfang im Nullvektor. Falls $y = (y_1, y_2)$ ein weiterer Vektor in der Ebene ist, so entspricht der Summe $x + y$ die Spitze des Parallelogramms, das von x und y aufgespannt wird. Die geometrische Interpretation im K^n ist analog.

(2) Der *Vektorraum der Folgen*: Sei K^∞ die Menge der Folgen in K, das heisst, die Menge der Folgen $x = (x_0, x_1, x_2, \ldots)$ mit $x_n \in K$ für alle $n \in \mathbb{N}$. Addition und skalare Multiplikation definieren wir wieder komponentenweise; wie im vorigen Beispiel setzen wir

$$x + y := (x_0 + y_0, x_1 + y_1, x_2 + y_2, \ldots) \quad \text{und} \quad \alpha x := (\alpha x_0, \alpha x_1, \alpha x_2, \ldots).$$

(3) Der *Vektorraum der Funktionen*: Sei X eine Menge. Abbildungen von X nach K nennen wir auch *Funktionen*. Sei K^X die Menge solcher Funktionen. Für $\alpha \in K$ und $f, g \in K^X$ müssen $f + g$ und αf wieder Funktionen, also Abbildungen von X nach K sein. Wir erklären Addition und skalare Multiplikation *punktweise*, das entspricht den Definitionen in den beiden vorherigen Beispielen, wenn wir n-Tupel als Funktionen $\{1, \ldots, n\} \to K$ und Folgen als Funktionen $\mathbb{N} \to K$ interpretieren: Für alle $x \in X$ sei

$$(f + g)(x) := f(x) + g(x) \quad \text{und} \quad (\alpha f)(x) := \alpha f(x).$$

(4) Der *Vektorraum der formalen Potenzreihen*: Sei x eine *Unbestimmte*, das heisst, ein Symbol. Eine *formale Potenzreihe* in x mit Koeffizienten in K ist eine formale Summe

$$\sum_{n \geq 0} a_n x^n = a_0 + a_1 x + a_2 x^2 + a_3 x^3 + \ldots,$$

wobei die *Koeffizienten* $a_0, a_1, a_2, a_3, \ldots$ aus K sind und $x^0 := 1$, $x^1 := x$ gesetzt wird. Wir nennen a_n den Koeffizienten von x^n. Mit $K[[x]]$ bezeichnen wir die Menge aller formalen Potenzreihen in x mit Koeffizienten in K. Wichtige Beispiele aus der Analysis sind Exponential-, Sinus- und Kosinusreihe:

$$\exp x = \sum_{j \geq 0} \frac{1}{j!} x^j = 1 + x + \frac{1}{2!} x^2 + \frac{1}{3!} x^3 + \ldots$$

$$\sin x = \sum_{j \geq 0} \frac{(-1)^j}{(2j+1)!} x^{2j+1} = x - \frac{1}{3!} x^3 + \frac{1}{5!} x^5 - \ldots$$

$$\cos x = \sum_{j \geq 0} \frac{(-1)^j}{(2j)!} x^{2j} = 1 - \frac{1}{2!} x^2 + \frac{1}{4!} x^4 - \ldots$$

An dieser Darstellung sehen wir mehreres: Nicht alle Potenzen von x müssen explizit auftreten, die Koeffizienten der nicht auftretenden Potenzen sind 0 und werden deshalb weggelassen. Wir setzen $x^0 := 1$, wir schreiben einfach x statt x^1 und x^n statt $1x^n$. Diese Konventionen sind praktisch, damit bleiben wir im Rahmen der üblichen Konventionen.

Addition und skalare Multiplikation in $K[[x]]$ definieren wir *koeffizientenweise*:

$$\alpha \sum_{n\geq 0} a_n x^n = \sum_{n\geq 0} (\alpha a_n) x^n,$$
$$\sum_{n\geq 0} a_n x^n + \sum_{n\geq 0} b_n x^n = \sum_{n\geq 0} (a_n + b_n) x^n.$$

In anderer Schreibweise:

$$\alpha \cdot (a_0 + a_1 x + a_2 x^2 + \ldots) = (\alpha a_0) + (\alpha a_1) x + (\alpha a_2) x^2 + \ldots$$

beziehungsweise

$$(a_0 + a_1 x + a_2 x^2 + \ldots) + (b_0 + b_1 x + b_2 x^2 + \ldots)$$
$$= (a_0 + b_0) + (a_1 + b_1) x + (a_2 + b_2) x^2 + \ldots$$

Mutige Geister sollten an dieser Stelle die Eulersche Formel $\exp(ix) = \cos x + i \sin x$ in $\mathbb{C}[[x]]$ ableiten.

Es gibt auch ein natürliches Produkt auf $K[[x]]$, dazu komme ich vielleicht später.

Im Folgenden bezeichnet V einen Vektorraum über K. Skalare bezeichnen wir zumeist mit griechischen Buchstaben, also mit $\alpha, \beta, \ldots$, Vektoren mit lateinischen, $v, w, \ldots$ Bevor wir zur nächsten Definition schreiten, einige kleine Bemerkungen:

(1) Für alle $\alpha \in K$ und $v \in V$ ist $\alpha \cdot 0 = 0$ und $0 \cdot v = 0$. Es gilt nämlich $\alpha \cdot 0 = \alpha \cdot (0 + 0) = \alpha \cdot 0 + \alpha \cdot 0$.

(2) Für alle $\alpha \in K$ und $v \in V$ gilt $\alpha v = 0 \Leftrightarrow \alpha = 0$ oder $v = 0$. Falls nämlich $\alpha \neq 0$ und $v \neq 0$, so ist $0 \neq v = 1 \cdot v = (\alpha^{-1}\alpha) v = \alpha^{-1}(\alpha v)$, also $\alpha v \neq 0$ nach (1). Nun folgt (2) aus (1).

(3) Für alle $\alpha \in K$ und $v \in V$ gilt $(-\alpha) \cdot v = \alpha \cdot (-v) = -(\alpha \cdot v)$.

8.3. DEFINITION. Sei V ein Vektorraum über K. Wir nennen eine Teilmenge U von V einen *Untervektorraum* oder *Unterraum* von V, falls

(1) $0 \in U$ und

(2) mit $\alpha \in K$ und $v, w \in U$ auch αv und $v + w$ in U sind.

Damit wird ein Unterraum zusammen mit den ererbten Verknüpfungen selber zum Vektorraum über K. Ein Unterraum enthält immer den Nullvektor, also ist der Durchschnitt von Unterräumen in V niemals leer.

8.4. BEISPIELE. Wir diskutieren Beispiele von Untervektorräumen U eines Vektorraumes V über K:

(0) Die beiden Extremfälle sind $U = V$ und der *triviale Unterraum* $U = \{0\}$.

(1) Lösungsmengen von Gleichungen: Diese Klasse von Beispielen werden wir später noch detaillierter besprechen, hier diskutieren wir nur ein erstes Beispiel. Seien $a, b, c \in K$ vorgegebene Skalare. Die Gleichung $ax + by = c$ heisst *linear* in den Unbekannten x und y, weil sie keine Terme quadratischer oder höherer Ordnung in den Unbekannten x und y enthält, wie zum Beispiel x^2, xy, oder y^2, sondern nur Terme

höchstens erster Ordnung in x und y. Die Gleichung heisst *inhomogen*, falls $c \neq 0$, *homogen*, falls $c = 0$ ist. Wir rechnen schnell nach, dass die Lösungsmenge

$$U := \{(x, y) \in K^2 \mid ax + by = c\}$$

genau dann ein Unterraum des K^2 ist, wenn $c = 0$, also die Gleichung homogen ist. Es gibt dann zwei Fälle: Falls $a = b = 0$ ist, dann ist $U = K^2$. Falls $(a, b) \neq (0, 0)$, so heisst die Gleichung *Geradengleichung*. Dann ist U ein echter Unterraum von K^2, $(-b, a)$ ist eine der Lösungen und jede andere Lösung ist ein skalares Vielfaches dieser Lösung.

(2) Im Raum V der Folgen in $\mathbb{R}$ oder $\mathbb{C}$ können wir die Teilmenge U der konvergenten Teilfolgen betrachten. Nach Sätzen aus der Analysis ist U ein Unterraum. Die Teilmenge U_0 der Nullfolgen ist ebenfalls ein Unterraum, es gilt $U_0 \subset U \subset V$.

(3) Ein Polynom in der Unbekannten x mit Koeffizienten in K ist eine formale Summe der Form

$$a_0 + a_1 x + a_2 x^2 + \ldots + a_n x^n,$$

wobei $n \geq 0$ ist. Das größte i mit $a_i \neq 0$ nennen wir den *Grad* des Polynoms. Der Grad des *Nullpolynoms*, also des Polynoms mit $a_i = 0$ für alle i, ist per definitionem gleich $-\infty$.

Wir können uns ein Polynom als formale Potenzreihe vorstellen, bei der die fehlenden Potenzen von x weggelassen wurden, weil ihre Koeffizienten 0 sind. In diesem Sinne ist die Menge $K[x]$ aller Polynome in der Unbekannten x mit Koeffizienten in K eine Teilmenge von $K[[x]]$. Wir rechnen schnell nach, dass $K[x]$ ein Unterraum ist. Insbesondere ist $K[x]$ mit der oben für $K[[x]]$ definierten Addition und skalaren Multiplikation selber ein Vektorraum über K.

Für jedes $n \in \mathbb{Z}$ ist die Menge der Polynome vom Grad $\leq n$ ein Unterraum von $K[x]$, damit auch von $K[[x]]$.

8.5. SATZ. *Seien U und U' Untervektorräume von V. Dann sind*

(1) *$U \cap U'$ und*

(2) Summe $U + U' := \{v + v' \mid v \in U, v' \in U'\}$

ebenfalls Untervektorräume von V.

Das analoge Resultat für mehr als zwei Untervektorräume gilt auch. Den Beweis überlassen wir als Übung.

9. Lineare Hülle, Dimension

Wir fixieren einen Vektorraum V über einem Körper K. Sei $E = (v_1, \dots, v_m)$ ein m-tupel von Vektoren in V. Eine Summe

$$\sum_{i=1}^{m} \alpha_i v_i = \alpha_1 v_1 + \ldots + \alpha_m v_m \tag{9.1}$$

mit $\alpha_i \in K$, $1 \le i \le m$, nennen wir eine *Linearkombination von* E. Wir lassen es dabei unbestimmt, ob wir damit die Summe insgesamt oder die Terme der Summe meinen. Der Nullvektor kann aus jedem Tupel auf triviale Weise linear kombiniert werden, nämlich als $0 = \sum 0 \cdot v_i$. Wenn $v \in V$ als Linearkombination von E dargestellt werden kann, sagen wir auch, dass v *linear von* E *abhängt*.

Die Menge $L(E)$ aller Linearkombinationen von E nennen wir die *lineare Hülle* oder den *Spann* von E. Umgekehrt nennen wir E ein *Erzeugendensystem* von $L(E)$.

9.2. Lemma. *Sei $E = (v_1, \dots, v_m)$ ein m-tupel von Vektoren in V. Dann ist*

(1) *$L(E)$ ein Unterraum von V.*
(2) *$v_i \in L(E)$, $1 \le i \le m$.*

Beweis. Seien $v = \sum \alpha_i v_i$, $w = \sum \beta_i v_i$ und $\gamma \in K$. Dann sind

$$v + w = \sum \alpha_i v_i + \sum \beta_i v_i = \sum (\alpha_i + \beta_i) v_i \in L(E),$$
$$\gamma v = \gamma \sum \alpha_i v_i = \sum (\gamma \alpha_i) v_i \in L(E),$$

damit (1). Sei $i \in \{1, \dots, m\}$. Setze $\alpha_j = 0$ für $j \ne i$ und $\alpha_i = 1$. Dann ist

$$v_i = 1 \cdot v_i + \sum_{j \ne i} 0 \cdot v_j = \sum \alpha_j v_j,$$

damit (2). □

9.3. Beispiel. Sei $V = K^n$, und für $i \in \{1, \dots, n\}$ sei e_i der $i-te$ *Einheitsvektor* in K^n, d.h., e_i ist der Vektor in K^n mit i-ter Komponente 1 und anderen Komponenten $= 0$. Für $m \le n$ besteht $L((e_1, \dots, e_m))$ dann aus allen Vektoren $x = (x_1, \dots, x_n) \in K^n$ mit $x_j = 0$ für $m < j \le n$. Insbesondere ist

$$L((e_1, \dots, e_n)) = K^n.$$

Es gibt ein kleines Problem mit der Notation. Es ist geschickter, bei den Summen ausser den Mengen $\{1, \dots, m\}$ beliebige endliche Indexmengen zuzulassen [13]. Sei dazu I zunächst eine beliebige Menge, die als *Indexmenge* fungiert. Die Vorgabe von Elementen x_i, $i \in I$, aus einer Menge X nennen wir ein *I-Tupel* in X, oft geschrieben als $(x_i)_{i \in I}$. Für $J \subset I$ eine (echte) Teilmenge nennen wir dann $(x_i)_{i \in J}$ ein *(echtes) Teiltupel*. Wenn es auf die Indexmenge nicht ankommt, dann sprechen wir auch einfach von Tupeln, notiert zum Beispiel als (x_i).

Im Folgenden werden die Indexmengen immer als endlich vorausgesetzt!

[13] Dies gilt allgemein.

Sei nun $E = (v_i)_{i\in I}$ ein I-Tupel von Vektoren in V. Eine Summe der Form

$$\sum_{i\in I} \alpha_i v_i\,, \tag{9.4}$$

mit $\alpha_i \in K$ für alle $i \in I$, nennen wir dann eine *Linearkombination* von E. Die Menge $L(E)$ aller Linearkombinationen von E nennen wir wieder die *lineare Hülle* oder den *Spann* von E.

9.5. LEMMA. *Sei $E = (v_i)_{i\in I}$ ein I-Tupel von Vektoren in V. Dann gilt:*

(1) *$L(E)$ ist ein Unterraum von V.*
(2) *Für alle $i \in I$ ist $v_i \in L(E)$.*
(3) *Falls U ein Unterraum von V ist mit $v_i \in U$ für alle $i \in I$, so ist $L(E) \subset U$.*
(4) *Falls F ein Teiltupel von E ist, so ist $L(F) \subset L(E)$.* □

9.6. LEMMA. *Sei $E = (v_i)_{i\in I}$ ein I-Tupel von Vektoren in V. Sei $j \in I$ ein Index, so dass v_j aus den v_i, $i \neq j$, linear kombiniert werden kann, $v_j = \sum_{i\neq j} \alpha_i v_i$. Dann ist $L((v_i)_{i\neq j}) = L(E)$.*

Beweis. Sei $v_j = \sum_{i\neq j} \alpha_i v_i$. Dies können wir in Linearkombinationen substituieren:

$$\sum_{i\in I} \beta_i v_i = \sum_{i\neq j} \beta_i v_i + \beta_j v_j = \sum_{i\neq j} \beta_i v_i + \beta_j \sum_{i\neq j} \alpha_i v_i = \sum_{i\neq j} (\beta_i + \beta_j \alpha_i) v_i.$$

Damit $L((v_i)_{i\neq j}) = L(E)$. □

Wir nennen E *linear abhängig*, wenn es ein $v \in L(E)$ gibt, das auf verschiedene Weisen aus E linear kombiniert werden kann, also

$$v = \sum_{i\in I} \alpha_i v_i = \sum_{i\in I} \beta_i v_i \tag{9.7}$$

mit $\alpha_i \neq \beta_i$ für zumindest ein $i \in I$. Falls E nicht linear abhängig ist, so nennen wir E *linear unabhängig*.

9.8. LEMMA. *Folgende Eigenschaften sind äquivalent:*

(1) *E ist linear abhängig.*
(2) *Der Nullvektor kann aus E nichttrivial linear kombiniert werden.*
(3) *Es gibt ein $j \in I$, so dass v_j von den v_i, $i \neq j$, linear abhängt.*
(4) *Es gibt eine echte Teilmenge $J \subset I$ mit $L((v_i)_{i\in J}) = L(E)$.*

Beweis. (1) ⇒ (2): Es gibt Linearkombinationen $\sum \alpha_i v_i = \sum \beta_i v_i$ mit $\alpha_j \neq \beta_j$ für ein $j \in I$. Dann ist

$$\sum_{i\in I} (\alpha_i - \beta_i) v_i = \sum_{i\in I} \alpha_i v_i - \sum_{i\in I} \beta_i v_i = 0,$$

aber $\alpha_j - \beta_j \neq 0$. Damit ist der Nullvektor nichttrivial linear kombiniert.
(2) ⇒ (3): Sei $\sum \alpha_i v_i = 0$, so dass $\alpha_j \neq 0$ für ein $j \in I$. Dann ist

$$v_j = -\frac{1}{\alpha_j} \sum_{i\neq j} \alpha_i v_i = \sum_{i\neq j} -\frac{\alpha_i}{\alpha_j} v_i.$$

(3) $\Rightarrow$ (4): Lemma 9.6 und $J = I \setminus \{j\}$.
(4) $\Rightarrow$ (1): Sei $J \subsetneqq I$ mit $L((v_i)_{i\in J}) = L(E)$. Sei $j \in I \setminus J$. Dann ist $v_j \in L(E) = L((v_i)_{i\in J})$, kann also aus den $v_i, i \in J$, linear kombiniert werden, $v_j = \sum_{i\in J} \alpha_i v_i$. Andererseits ist $v_j = \sum \beta_i v_i$ mit $\beta_j = 1$ und $\beta_i = 0$ für $i \neq j$. Damit kann v_j auf zwei verschiedene Weisen aus E linear kombiniert werden. □

9.9. LEMMA. *Sei $F = (w_i)_{i\in I}$ ein linear unabhängiges I-Tupel in V, und sei $v \in V \setminus L(F)$. Sei J Obermenge von I mit $J \setminus I = \{j\}$. Definiere ein J-Tupel $E = (v_i)_{i\in J}$ durch $v_i := w_i$ für $i \neq j$ und $v_j := v$. Dann ist F Teiltupel von E und E ist ebenfalls linear unabhängig.*

Beweis. Sei $\sum_{i\in J} \alpha_i v_i = 0$. Dann ist

$$-\alpha_j v_j = \sum_{i\in I} \alpha_i w_i \in L(F).$$

Nun ist $v_j = v \notin L(F)$, also ist $\alpha_j = 0$ und damit $\sum_{i\in I} \alpha_i w_i = 0$. Weil F linear unabhängig ist, folgt $\alpha_i = 0$ für alle $i \in I$. □

9.10. DEFINITION. Ein I-Tupel E in V heißt *Basis* von V, wenn E linear unabhängig und $L(E) = V$ ist.
9.11. SATZ. *Sei E ein I-Tupel in V. Dann sind äquivalent:*

(1) *E ist Basis von V.*
(2) *E ist minimales Erzeugendensystem von V, d.h., E enthält kein echtes Teiltupel E' mit $L(E') = V$.*
(3) *E ist maximales linear unabhängiges Tupel in V, d.h., E ist kein echtes Teiltupel eines linear unabhängigen Tupels in V.*

Beweis. (1) $\Rightarrow$ (2): Falls E ein echtes Teiltupel E' mit $L(E') = V$ enthalten würde, so wäre E nach Lemma 9.8 linear abhängig.
(2) $\Rightarrow$ (3): Falls E echtes Teiltupel eines linear unabhängigen Tupels E' in V wäre, so wäre $L(E) \subset L(E')$ nach Lemma 9.8 ein echter Unterraum. Wegen $L(E') \subset V$ wäre dann aber $L(E) \neq V$.
(3) $\Rightarrow$ (1): Wäre $L(E) \neq V$, so gäbe es ein $v \in V \setminus L(E)$. Dann könnten wir eine Obermenge J von I wählen mit $J \setminus I = \{j\}$. Mit $v_j := v$ wäre dann die Erweiterung $(v_i)_{i\in J}$ von E nach Lemma 9.9 linear unabhängig. □

Wir sagen, dass V *endliche Dimension* hat und schreiben $\dim V < \infty$, wenn es ein I-Tupel $(v_i)_{i\in I}$ mit $|I| < \infty$ gibt, so dass $L(E) = V$ ist. Andernfalls sagen wir, dass V *unendliche Dimension* hat, $\dim V = \infty$.
9.12. SATZ. *Sei $|I| < \infty$ und $E = (v_i)_{i\in I}$ ein I-Tupel in V mit $L(E) = V$. Sei $K \subset I$ eine Teilmenge, so dass $F = (v_i)_{i\in K}$ linear unabhängig ist. Dann gibt es J zwischen K und I, also $K \subset J \subset I$, so dass $B = (v_i)_{i\in J}$ eine Basis von V ist.*

Beweis. Unter den Teilmengen von I, die K enthalten, sei J maximal mit der Eigenschaft, dass $B = (v_i)_{i\in J}$ linear unabhängig ist. Falls $L(B) \neq V$ wäre, so gäbe es ein

$j \in I$ mit $v_j \notin L(B)$, denn $L(E) = V$. Aber dann wäre $(v_i)_{i\in J\cup\{j\}}$ nach Lemma 9.9 linear unabhängig. Das steht im Widerspruch zur Wahl von J. □

9.13. KOROLLAR. *Sei* $\dim V < \infty$. *Dann gilt:*

(1) V *hat eine Basis.*
(2) *Jedes endliche Erzeugendensystem von* V *enthält eine Basis.*
(3) *Jedes linear unabhängige Tupel kann zu einer Basis ergänzt werden.* □

9.14. LEMMA. *Sei* $\dim V < \infty$ *und seien* $B = (v_i)_{i\in I}$ *und* $C = (w_j)_{j\in J}$ *zwei Basen von* V. *Dann gibt es zu jedem* $i_0 \in I$ *ein* $j_0 \in J$, *so dass das* I*-Tupel* $B' = (v'_i)_{i\in I}$ *mit* $v'_i := v_i$ *für* $i \neq i_0$ *und* $v'_{i_0} = w_{j_0}$ *wieder eine Basis von* V *ist.*

Beweis. Weil B minimales Erzeugendensystem von V ist, folgt $L((v_i)_{i\neq i_0}) \neq V$. Nun ist $L(C) = V$, daher gibt es ein $j_0 \in J$ mit $w_{j_0} \notin L((v_i)_{i\neq i_0})$. Mit diesem j_0 ist dann B' nach Lemma 9.9 linear unabhängig. Wäre $L(B') \neq V$, so wäre $v_{i_0} \notin L(B')$, denn $L(B) = V$. Dann wäre aber B' zusammen mit v_{i_0} nach Lemma 9.9 linear unabhängig. Dieses neue Tupel wäre nun umgekehrt eine echte Erweiterung von B, ein Widerspruch zur Maximalität von B als linear unabhängiges Tupel. □

9.15. SATZ. *Sei* $\dim V < \infty$ *und seien* $B = (v_i)_{i\in I}$ *und* $C = (w_j)_{j\in J}$ *zwei Basen von* V. *Dann ist* $|I| = |J|$.

Die Zahl $|I|$ nennen wir die *Dimension* von V.

Beweis des Satzes. Sei o.B.d.A. $|I| \leq |J|$. Mit dem Lemma oben können wir sukzessive die Vektoren v_i, $i \in I$, durch Vektoren $w_{j(i)}$, $j(i) \in J$, ersetzen und erhalten eine neue Basis $B' = (w_{j(i)})_{i\in I}$. Offenbar ist die lineare Hülle von B' gleich der linearen Hülle des Teiltupels von C, das aus allen in B' vorkommenden w_j besteht. Nun ist $L(B') = V$, denn B' ist eine Basis. Also sind alle j aus J ein $j(i)$, damit $|J| \geq |I|$. □

9.16. BEISPIELE. 1) Seien $e_1, \ldots, e_n$ die Einheitsvektoren in K^n. Sei $x = (x_1, \ldots, x_n)$ ein Vektor in K^n. Dann ist

$$x = \sum x_i e_i.$$

Also ist $E = (e_1, \ldots, e_n)$ ein Erzeugendensysten des K^n. Nun trägt e_i bei Linearkombinationen nur zur i-ten Komponente bei, daher ist E auch linear unabhängig und damit eine Basis von K^n. Also ist $\dim K^n = n$.

2) Im Raum K^∞ der Folgen sei $e_i = (x_0, x_1, x_2, \ldots)$ die Folge mit $x_i = 1$ und $x_j = 0$ für $j \neq i$. Wie im ersten Beispiel folgt, dass jedes endliche Teiltupel von $(e_1, e_2, \ldots)$ linear unabhängig ist. Also ist $\dim K^\infty = \infty$.

9.17. KOROLLAR. *Sei* $\dim V < \infty$ *und sei* $E = (v_i)$ *ein* I*-Tupel von Vektoren in* V. *Dann gelten die beiden (äquivalenten) Aussagen:*

(1) $|I| > \dim V \Rightarrow E$ *ist linear abhängig.*
(2) E *ist linear unabhängig* $\Rightarrow |I| \leq \dim V$.

Beweis. Klar mit Korollar 9.13 und Satz 9.15. □

Wir zeigen noch eine Erweiterung von Lemma 9.14.

9.18. SATZ (Austauschsatz). [14] *Sei* $\dim V < \infty$, $B = (v_i)_{i\in I}$ *eine Basis von* V *und* $F = (w_i)_{i\in J}$ *ein linear unabhängiges Tupel. Dann gibt es eine Teilmenge* $K \subset I$ *mit* $|K| = |I| - |J|$, *so dass das* $(K \sqcup J)$*-Tupel* [15] $B' = (v'_i)$ *mit*

$$v'_i := \begin{cases} v_i & \text{für } i \in K, \\ w_i & \text{für } i \in J, \end{cases}$$

eine Basis von V *ist.*

Beweis. Setze $I' = I \sqcup J$. Dann ist das I'-Tupel $E' = (v'_i)$ mit $v'_i := v_i$ für $i \in I$ und $v'_i := w_i$ für $i \in J$ ein Erzeugendensystem von V. Nun ist F linear unabhängig und Teiltupel von E'. Nach Satz 9.12 gibt es daher eine Teilmenge $K \subset I$, so dass das entsprechende $(K \sqcup J)$-Teiltupel eine Basis von V ist. Wegen Satz 9.15 muss $|K| = |I| - |J|$ sein. □

9.19. KOROLLAR. *Sei* V *endlichdimensional und* U *ein Unterraum von* V. *Dann ist auch* U *endlichdimensional und* $\dim U \le \dim V$. *Jede Basis von* U *kann zu einer Basis von* V *erweitert werden.*

Beweis. Weil linear unabhängige Tupel in U auch linear unabhängig in V sind, bestehen solche aus höchstens $\dim V$ Mitgliedern. Wenn E ein solches mit einer maximalen Anzahl von Mitgliedern ist, so muss E nach Satz 9.11 eine Basis von U sein. Also ist $\dim U \le \dim V$. Die letzte Behauptung folgt sofort aus Satz 9.12. □

9.20. SATZ (Dimensionsformel für Unterräume). *Seien* U *und* U' *endlichdimensionale Unterräume von* V. *Dann sind auch* $U \cap U'$ *und* $U + U'$ *endlichdimensional und*

$$\dim(U + U') + \dim(U \cap U') = \dim U + \dim U'.$$

Genauer gilt: Falls $E = (v_1, \dots, v_r)$ *eine Basis von* $U \cap U'$ *ist und*

$$B = (v_1, \dots, v_r, w_1, \dots, w_s) \quad \textit{und} \quad B' = (v_1, \dots, v_r, w'_1, \dots, w'_t)$$

Erweiterungen von E *zu Basen von* U *und* U' *sind, so ist*

$$C = (v_1, \dots, v_r, w_1, \dots, w_s, w'_1, \dots, w'_t)$$

eine Basis von $U + U'$.

Beweis. Offenbar ist C ein Erzeugendensystem von $U + U'$, also $\dim(U + U') < \infty$. Sei nun

$$\sum \alpha_i v_i + \sum \beta_j w_j + \sum \beta'_k w'_k = 0.$$

Dann ist

$$\sum \beta_j w_j = -\sum \alpha_i v_i - \sum \beta'_k w'_k$$

in U und U', also in $U \cap U' = L(E)$. Dann kann die linke Seite aber aus E linear kombiniert werden. Nun ist B linear unabhängig, also sind die $\beta_j = 0$. Weil

[14] Nach Ernst Steinitz (1871–1928) auch *Steinitzscher Austauschsatz* genannt. Siehe hierzu die historischen Bemerkungen in [Jä].

[15] $\sqcup$ bezeichnet die disjunkte Vereinigung.

B' linear unabhängig ist, müssen deshalb auch die übrigen Koeffizienten α_i und β'_k verschwinden. Also ist C linear unabhängig. □

9.21. KOROLLAR. *Sei* $\dim V < \infty$*, und seien* U *und* U' *Unterräume von* V *mit* $U \cap U' = \{0\}$ *und* $U + U' = V$*. Dann ist*

$$\dim V = \dim U + \dim U'.$$

9.22. BEMERKUNG. Sei $\dim V < \infty$ und $U \subset V$ ein Unterraum. Dann gibt es einen Unterraum U' von V mit $U \cap U' = \{0\}$ und $U + U' = V$. Sei nämlich $(v_1, \ldots, v_m)$ eine Basis von U. Ergänze diese zu einer Basis $(v_1, \ldots, v_m, w_1, \ldots, w_n)$ von V. Setze $U' = L((w_1, \ldots, w_n))$.

Wir diskutieren jetzt noch eine Anwendung unserer Ergebnisse auf die Theorie linearer Gleichungen. Seien $\alpha_1, \ldots, \alpha_n, \beta \in K$. Dann nennt man

$$\alpha_1 x_1 + \ldots + \alpha_n x_n = \beta$$

eine *lineare Gleichung* in den Unbekannten $x_1, \ldots, x_n$. Gesucht sind Lösungen $x = (x_1, \ldots, x_n) \in K^n$. Allgemeiner sucht man nach simultanen Lösungen von m Gleichungen dieser Art,

$$\begin{array}{c} \alpha_{11}x_1 + \alpha_{12}x_2 + \ldots + \alpha_{1n}x_n = \beta_1, \\ \alpha_{21}x_1 + \alpha_{22}x_2 + \ldots + \alpha_{2n}x_n = \beta_2, \\ \vdots \qquad \vdots \qquad\qquad \vdots \quad \vdots \\ \alpha_{m1}x_1 + \alpha_{m2}x_2 + \ldots + \alpha_{mn}x_n = \beta_m. \end{array} \tag{9.23}$$

Wir setzen nun

$$a_1 := \begin{pmatrix} \alpha_{11} \\ \vdots \\ \alpha_{m1} \end{pmatrix}, \ldots, a_n := \begin{pmatrix} \alpha_{1n} \\ \vdots \\ \alpha_{mn} \end{pmatrix}, \; b := \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_m \end{pmatrix} \in K^m.$$

Dann können wir (9.23) schreiben als

$$\sum_{i=1}^{n} x_i a_i = x_1 a_1 + \ldots + x_n a_n = b. \tag{9.24}$$

Daher ist (9.23) genau dann lösbar, wenn $b \in L((a_1, \ldots, a_n))$. Wir definieren noch

$$\operatorname{Rang}(a_1, \ldots, a_n) := \dim L((a_1, \ldots, a_n)). \tag{9.25}$$

Damit können wir das Ergebnis unserer Diskussion wie folgt zusammenfassen.

9.26. SATZ. *Die folgenden Aussagen sind äquivalent:*

(1) *Das Gleichungssystem 9.23 ist lösbar.*
(2) $b \in L((a_1, \ldots, a_n))$.
(3) $\operatorname{Rang}(a_1, \ldots, a_n) = \operatorname{Rang}(a_1, \ldots, a_n, b)$. □

Zum Schluss noch eine Erweiterung der Definition von *linear unabhängig* und *linearer Hülle eines Tupels* $E = (v_i)_{i\in I}$ *in* V auf den Fall, dass I unendlich ist: Wir sagen, dass E linear unabhängig ist, wenn für jede endliche Teilmenge $J \subset I$ das Teiltupel $(v_i)_{i\in J}$ von E linear unabhängig ist. Wir sagen, dass $v \in V$ linear von E abhängt, wenn es eine endliche Teilmenge $J \subset I$ gibt, so dass v linear abhängig von dem Teiltupel $(v_i)_{i\in J}$ ist. Die lineare Hülle $L(E)$ ist dann die Menge aller von E linear abhängigen Vektoren aus V. Eine Basis eines Vektorraumes V über K, ob endlich- oder unendlichdimensional, ist dann ein linear unabhängiges Tupel E mit $L(E) = V$.

Wir haben bewiesen, dass jeder endlichdimensionale Vektorraum eine Basis besitzt, aber dies gilt auch für beliebige Vektorräume. Im Beweis benützt man auf die eine oder andere Weise das Auswahlaxiom beziehungsweise das dazu äquivalente Lemma von Zorn. Wir werden dieses allgemeinere Ergebnis weder beweisen noch benützen.

9.27. BEMERKUNG. Das in Beispiel 9.16 definierte Tupel von Vektoren $(e_i)_{i\geq 0}$ in K^∞ ist linear unabhängig, aber keine Basis von K^∞. Zum Beispiel kann der Vektor $(1, 1, 1, \ldots)$ nicht aus diesen Vektoren linear kombiniert werden, denn nach Definition sind nur endliche Linearkombinationen der e_i zulässig.

10. Lineare Abbildungen und Matrizen

10.1. Definition. Seien V, W Vektorräume über K und $f : V \to W$ eine Abbildung. Wir sagen, dass f *linear* ist, wenn für alle $v, v' \in V$ und $\alpha \in K$ gilt

$$\begin{aligned} f(\alpha v) &= \alpha f(v) && (\textit{Homogenität}),\\ f(v + v') &= f(v) + f(v') && (\textit{Additivität}). \end{aligned}$$

Die Menge aller linearen Abbildungen von V nach W bezeichnen wir mit $\mathrm{Hom}(V, W)$.

Die Abkürzung Hom bezieht sich auf einen anderen gebräuchlichen Namen für lineare Abbildungen, sie heissen auch *Homomorphismen von Vektorräumen* (über K). Dies bedeutet übersetzt, dass sie mit den auf den Vektorräumen vorgegebenen Verknüpfungen, Addition von Vektoren und Multiplikation mit Skalaren, verträglich sind, und zwar genau im Sinne von Definition 10.1. Speziell nennt man eine lineare Abbildung $f : V \to W$ einen

(1) *Endomorphismus*, wenn $V = W$ ist.
(2) *Isomorphismus*, wenn f bijektiv ist.
(3) *Automorphismus*, wenn $V = W$ und f bijektiv ist.

Entsprechend definiert man Homo-, Endo-, Iso- und Automorphismen von Gruppen, Ringen usw. über die Verträglichkeit der Abbildungen mit *allen* Verknüpfungen, durch die die jeweilige Struktur gegeben ist. Es lohnt daher, sich diese Begriffe einzuprägen.

Während in unserer bisherigen Diskussion die Vektorräume fest blieben, kommt mit den linearen Abbildungen eine gewisse Dynamik in's Spiel: Unter der linearen Abbildung $f : V \to W$ wird den Vektoren aus V jeweils ein Vektor aus W zugeordnet, die Vektoren *wandern von V nach W*.

10.2. Beispiele. 0) Die *triviale Abbildung* oder *Nullabbildung* $f : V \to W$, $f(v) := 0$ für alle $v \in V$, ist linear.

1) Die identische Abbildung $\mathrm{id}_V : V \to V$ ist linear, denn

$$\mathrm{id}_V(v + v') = v + v' = \mathrm{id}_V(v) + \mathrm{id}_V(v'), \quad \mathrm{id}_V(\alpha v) = \alpha v = \alpha\, \mathrm{id}_V(v).$$

Sei allgemeiner $U \subset V$ ein Unterraum. Dann ist die *Inklusion* $i : U \to V$, $i(v) := v$, linear.

2) Sei $V = K^n$, und $p_i : V \to K$ die *Projektion* auf die i-te Komponente, $p_i(x) := x_i$ für $x = (x_1, \ldots, x_n)$. Hier betrachten wir $K = K^1 =: W$ als (eindimensionalen) Vektorraum über sich selber.

3) Sei K^∞ der Raum der Folgen in K und $S : K^\infty \to K^\infty$ der *Shift*,

$$S((x_0, x_1, x_2, \ldots)) = (x_1, x_2, x_3, \ldots),$$

d.h., wir streichen das erste Glied der Folge. Dann ist S (surjektiv und) linear.

4) Sei $V = W$, und seien $\lambda \in K$ und $w \in V$. Die *Streckung* oder *Homothetie* $S_\lambda : V \to V$, $v \mapsto \lambda v$, ist ein Endomorphismus von V. Die *Translation* $T_w : V \to V$, $v \mapsto v + w$, ist genau dann linear, wenn $w = 0$, also $T_w = \mathrm{id}_V$ ist.

Wir diskutieren jetzt die wichtigste Beispielklasse. Eine $(m\times n)$-*Matrix* mit *Einträgen* oder *Komponenten* in K ist ein rechteckiges Schema

$$A = \begin{pmatrix} \alpha_{11} & \alpha_{12} & \dots & \alpha_{1n} \\ \alpha_{21} & \alpha_{22} & \dots & \alpha_{2n} \\ \vdots & \vdots & & \vdots \\ \alpha_{m1} & \alpha_{m2} & \dots & \alpha_{mn} \end{pmatrix} \tag{10.3}$$

mit $\alpha_{ij} \in K, 1 \le i \le m, 1 \le j \le n$. Zahlenschemata dieser Art sind uns schon bei den linearen Gleichungssystemen im letzten Abschnitt begegnet. Wir wollen nun einer solchen Matrix eine lineare Abbildung $K^n \to K^m$ zuordnen. Für den Moment taufen wir diese f_A, später werden wir sie dann einfachheitshalber mit A bezeichnen [16]: Wir interpretieren die linke Seite des Gleichungssystems (9.23) als Zuordnungsvorschrift. Mit anderen Worten, $f_A(x) =: y = (y_1, \dots, y_m) \in K^m$ ist gegeben durch

$$\begin{aligned} y_1 &= \alpha_{11}x_1 + \alpha_{12}x_2 + \dots + \alpha_{1n}x_n, \\ y_2 &= \alpha_{21}x_1 + \alpha_{22}x_2 + \dots + \alpha_{2n}x_n, \\ &\vdots \qquad\quad \vdots \qquad\quad \vdots \qquad\qquad \vdots \\ y_m &= \alpha_{m1}x_1 + \alpha_{m2}x_2 + \dots + \alpha_{mn}x_n, \end{aligned} \tag{10.4}$$

wobei die Komponenten von y vertikal untereinander angeordnet sind. Wir beschreiben jetzt ein Schema, das sehr einprägsam ist und die Berechnungsvorschrift darstellt: Wir stellen uns x und y senkrecht geschrieben vor. [17] Für jedes $i \in \{1, \dots, m\}$ legen wir den Vektor x horizontal auf die i-te Zeile von A, multiplizieren die sich dabei entsprechenden Elemente α_{ij} und x_j und bilden die Summe über diese Produkte,

$$y_i = \sum \alpha_{ij}x_j = \alpha_{i1}x_1 + \dots + a_{in}x_n. \tag{10.5}$$

Nun haben wir die Zuordnungsvorschrift. Es bleibt zu zeigen, dass f_A linear ist. Seien dazu $x = (x_1, \dots, x_n)$ und $x' = (x'_1, \dots, x'_n)$ in K^n. Dann ist

$$x + x' = (x_1 + x'_1, \dots, x_n + x'_n)$$

und damit

$$\begin{aligned} f_A(x + x') &= \big(\sum \alpha_{1j}(x_j + x'_j), \dots, \sum \alpha_{mj}(x_j + x'_j)\big) \\ &= \big(\sum \alpha_{1j}x_j + \sum \alpha_{1j}x'_j, \dots, \sum \alpha_{mj}x_j + \sum \alpha_{mj}x'_j\big) \\ &= \big(\sum \alpha_{1j}x_j, \dots, \sum \alpha_{mj}x_j\big) + \big(\sum \alpha_{1j}x'_j, \dots, \sum \alpha_{mj}x'_j\big) \\ &= f_A(x) + f_A(x'), \end{aligned}$$

also ist f_A additiv. Die Homogenität von f_A folgt mit einer analogen Rechnung direkt daraus, dass K ein Ring ist. Das Resultat unserer Diskussion: Einer $(m \times n)$-Matrix

[16] Wir sprechen dann von der $(m \times n)$-Matrix, *aufgefasst als lineare Abbildung* $K^n \to K^m$.

[17] Aus Platzgründen bleiben wir aber im Text in der Regel bei der horizontalen Anordnung der Koeffizienten.

A mit Einträgen in K ordnen wir durch (10.5) eine lineare Abbildung $f_A : K^n \to K^m$ zu. Die Menge aller $(m \times n)$-Matrizen mit Einträgen in K bezeichnen wir mit $\mathrm{Mat}(m \times n, K)$. Damit erhalten wir eine Abbildung

$$\mathrm{Mat}(m \times n, K) \to \mathrm{Hom}(K^n, K^m), \quad A \mapsto f_A. \tag{10.6}$$

Weiter unten werden wir sehen, dass diese Abbildung bijektiv ist.

10.7. LEMMA. *Sei $\dim V < \infty$ und $B = (v_i)_{i \in I}$ eine Basis von V. Sei W ein weiterer Vektorraum über K. Dann gibt es zu jedem beliebigen I-Tupel $C = (w_i)_{i \in I}$ in W eine und genau eine lineare Abbildung $f : V \to W$ mit $f(v_i) = w_i$ für alle $i \in I$. Diese ist*

(1) *injektiv genau dann, wenn C linear unabhängig ist.*
(2) *surjektiv genau dann, wenn C ein Erzeugendensystem von W ist.*
(3) *bijektiv genau dann, wenn C eine Basis von W ist.*

Beweis. Sei $v \in V$, $v = \sum \alpha_i v_i$. Setze $f(v) := \sum \alpha_i w_i$. Da v auf genau eine Weise aus B linear kombiniert werden kann, ist f wohldefiniert und erfüllt $f(v_i) = w_i$ wie gewünscht. Seien $v, v' \in V$, $v = \sum \alpha_i v_i$ und $v' = \sum \alpha_i' v_i$. Dann ist $v + v' = \sum(\alpha_i + \alpha_i')v_i$ und damit

$$f(v + v') = \sum(\alpha_i + \alpha_i')w_i = \sum \alpha_i w_i + \sum \alpha_i' w_i = f(v) + f(v').$$

Also ist f additiv. Analog sieht man, dass f homogen ist. Umgekehrt sieht man an diesen Rechnungen, dass die Abbildung f durch ihre Werte auf B bestimmt ist.

Falls C linear abhängig ist, dann gibt es einen Vektor $w \in L(C)$, der auf zwei verschiedene Weisen aus C linear kombiniert werden kann, $w = \sum \alpha_i w_i = \sum \beta_i w_i$ mit $\alpha_i \neq \beta_i$ für zumindest ein $i \in I$. Die Vektoren $v = \sum \alpha_i v_i$ und $v' = \sum \beta_i v_i$ in V sind dann verschieden, denn B ist eine Basis von V, haben aber unter f dasselbe Bild, nämlich w. Daher ist f nicht injektiv, falls C linear abhängig ist. Entsprechend folgen die restlichen Behauptungen. □

10.8. KOROLLAR. *Endlichdimensionale Vektorräume über K sind genau dann isomorph, wenn sie die gleiche Dimension haben. Genauer gilt: Falls V und W Vektorräume über K mit $\dim V = \dim W < \infty$ sind und $B = (v_1, \ldots, v_n)$ und $C = (w_1, \ldots, w_n)$ Basen von V und W, so ist die eindeutig bestimmte lineare Abbildung $f : V \to W$ mit $f(v_i) = w_i$ ein Isomorphismus.* □

10.9. KOROLLAR. *Falls V und W Vektorräume über K mit $\dim V = \dim W < \infty$ sind und $f : V \to W$ eine lineare Abbildung ist, so ist*

$$f \text{ injektiv} \iff f \text{ surjektiv} \iff f \text{ bijektiv}. \quad \square$$

Zurück zu unserer wichtigsten Beispielklasse, linearen Abbildungen $K^n \to K^m$ und Matrizen. Mit Lemma 10.7 wissen wir, dass eine lineare Abbildung $f : K^n \to K^m$ eindeutig durch die Bilder $f(e_j)$ der Einheitsvektoren e_j, $1 \leq j \leq n$, festgelegt ist. Die $f(e_j)$ drücken wir als Linearkombinationen in der Standardbasis $C = (f_1, \ldots, f_m)$

des K^m aus,

$$f(e_j) = \sum_{i=1}^{m} \alpha_{ij} f_i = (\alpha_{1j}, \dots, \alpha_{mj}). \tag{10.10}$$

Die Koeffizienten α_{ij} ordnen wir in einer $(m \times n)$-Matrix an,

$$A := \begin{pmatrix} \alpha_{11} & \alpha_{12} & \dots & \alpha_{1n} \\ \alpha_{21} & \alpha_{22} & \dots & \alpha_{2n} \\ \vdots & \vdots & & \vdots \\ \alpha_{m1} & \alpha_{m2} & \dots & \alpha_{mn} \end{pmatrix}. \tag{10.11}$$

In der j-ten Spalte dieser Matrix steht also genau das Bild $f(e_j)$ von e_j, wobei wir die Komponenten von $f(e_j)$ untereinander, also senkrecht anordnen. Sei nun $x = (x_1, \dots, x_n)$ ein (beliebiger) Vektor in K^n. Dann ist $x = \sum x_j e_j$ nach Definition der Einheitsvektoren, damit

$$\begin{aligned} f(x) &= f(\sum_j x_j e_j) = \sum_i f(x_j e_j) \\ &= \sum_j x_j f(e_j) = \sum_{i,j} x_j \alpha_{ij} f_i = \sum_i \big(\sum_j \alpha_{ij} x_j\big) f_i \end{aligned}$$

wegen der Linearität von f. Die i-te Komponente von $f(x)$ ist daher $\sum_j \alpha_{ij} x_j$, genau wie bei der Abbildung f_A weiter oben. Daher ist $f = f_A$, wobei die Matrix A gegeben ist durch die Vektoren $f(e_1), \dots, f(e_n)$, hintereinander geschrieben als Spalten von A. Es folgt nun sofort, dass die Abbildung 10.6,

$$\mathrm{Mat}(m \times n, K) \to \mathrm{Hom}(K^n, K^m), \quad A \mapsto f_A,$$

bijektiv ist, denn sie ist ohne Zweifel auch injektiv.

10.12. LEMMA. *Seien $f, g : V \to W$ lineare Abbildungen und $\lambda \in K$. Dann sind die Abbildungen $\lambda f : V \to W$ und $(f + g) : V \to W$, definiert durch*

$$(\lambda f)(v) := \lambda f(v) \quad \textit{und} \quad (f + g)(v) := f(v) + g(v)$$

wieder linear. Mit diesen Verknüpfungen ist $\mathrm{Hom}(V, W)$ ein Vektorraum über K. Das neutrale Element der Addition ist die Nullabbildung.

Beweis. Seien $v \in V$ und $\alpha \in K$. Weil f linear ist, folgt

$$\begin{aligned} (\lambda f)(\alpha v) &= \lambda \cdot f(\alpha v) = \lambda \cdot (\alpha \cdot f(v)) \\ &= (\lambda\alpha) \cdot f(v) = (\alpha\lambda) \cdot f(v) \\ &= \alpha \cdot (\lambda \cdot f(v)) = \alpha \cdot (\lambda f)(v). \end{aligned}$$

Dies zeigt, dass λf homogen ist. Additivität von λf und Linearität von $f + g$ folgen mit ähnlichen Rechnungen. Wir überlassen es als Übung, zu zeigen, dass $\mathrm{Hom}(V, W)$ mit diesen Verknüpfungen ein Vektorraum über K ist. □

Die Menge $\mathrm{Mat}(m \times n, K)$ wird in der üblichen Weise zu einem Vektorraum über K: Eine Matrix A wie oben können wir interpretieren als die Abbildung $(i,j) \mapsto \alpha_{ij}$ definiert auf der Menge X der Paare (i,j) mit $1 \le i \le m$ und $1 \le j \le n$ und mit Werten in K. Damit haben wir auch schon das Rezept für die Verknüpfungen in $\mathrm{Mat}(m \times n, K)$: Für $A = (\alpha_{ij}), B = (\beta_{ij}) \in \mathrm{Mat}(m \times n, K)$ und $\lambda \in K$ setzen wir

$$A + B := \begin{pmatrix} \alpha_{11} + \beta_{11} & \alpha_{12} + \beta_{12} & \dots & \alpha_{1n} + \beta_{1n} \\ \alpha_{21} + \beta_{21} & \alpha_{22} + \beta_{22} & \dots & \alpha_{2n} + \beta_{2n} \\ \vdots & \vdots & & \vdots \\ \alpha_{m1} + \beta_{m1} & \alpha_{m2} + \beta_{m2} & \dots & \alpha_{mn} + \beta_{mn} \end{pmatrix}, \tag{10.13}$$

das heisst: wir addieren jeweils die Einträge von A und B mit gleichem Indexpaar. Analog verfahren wir bei der skalaren Multiplikation,

$$\lambda A := \begin{pmatrix} \lambda\alpha_{11} & \lambda\alpha_{12} & \dots & \lambda\alpha_{1n} \\ \lambda\alpha_{21} & \lambda\alpha_{22} & \dots & \lambda\alpha_{2n} \\ \vdots & \vdots & & \vdots \\ \lambda\alpha_{m1} & \lambda\alpha_{m2} & \dots & \lambda\alpha_{mn} \end{pmatrix}. \tag{10.14}$$

Wenn wir Matrizen als Abbildungen auf der Menge X der Paare wie oben ansehen, entsprechen diese Verknüpfungen genau den üblichen Verknüpfungen auf der Menge K^X. Es folgt, dass $\mathrm{Mat}(m \times n, K)$ zusammen mit diesen Verknüpfungen ein Vektorraum über K ist. Seine Dimension ist die Anzahl der Paare von Indizes (i,j), also $= m \cdot n$.

Offensichtlich entsprechen Addition und skalare Multiplikation in $\mathrm{Mat}(m \times n, K)$ genau Addition und skalarer Multiplikation im Raum der linearen Abbildungen $K^n \to K^m$; mit anderen Worten, die Abbildung 10.6,

$$\mathrm{Mat}(m \times n, K) \to \mathrm{Hom}(K^n, K^m), \quad A \mapsto f_A,$$

ist bezüglich dieser Verknüpfungen linear, also ein Isomorphismus von Vektorräumen. Damit folgt insbesondere $\dim \mathrm{Hom}(K^n, K^m) = m \cdot n$.

10.15. LEMMA. *Falls $f : U \to V$ und $g : V \to W$ linear sind, so ist auch die Komposition $g \circ f : U \to W$ linear.*

Beweis. Seien $v, v' \in U$. Dann ist

$$\begin{aligned}(g \circ f)(v + v') &= g(f(v + v')) = g(f(v) + f(v')) \\ &= g(f(v)) + g(f(v')) = (g \circ f)(v) + (g \circ f)(v'),\end{aligned}$$

also ist $g \circ f$ additiv. Die Homogenität folgt mit einer analogen Rechnung. □

Die Komposition von Abbildungen ist assoziativ, insbesondere ist die Komposition linearer Abbildungen assoziativ: Falls $f : U \to V$, $g : V \to W$ und $h : W \to X$ lineare Abbildungen sind, so ist $h \circ (g \circ f) = (h \circ g) \circ f$. Nach Lemma 10.15 ist diese Abbildung $U \to X$ linear.

10.16. LEMMA. *Falls $f : V \to W$ linear und bijektiv ist, so ist die Umkehrabbildung f^{-1} ebenfalls linear.*

Beweis. Seien $w, w' \in W$. Setze $v = f^{-1}(w)$, $v' = f^{-1}(w')$. Dann gilt

$$f(v) = w,\ f(v') = w',\ f(v + v') = f(v) + f(v') = w + w',$$

nach der Definition der Umkehrabbildung und weil f linear ist. Damit ist

$$f^{-1}(w + w') = v + v' = f^{-1}(w) + f^{-1}(w'),$$

also ist f additiv. Die Homogenität folgt mit einer analogen Rechnung. □

Die Teilmenge $\mathrm{Aut}(V) \subset \mathrm{End}(V) := \mathrm{Hom}(V, V)$ der Automorphismen von V wird damit zu einer Gruppe: Die Komposition linearer Abbildungen ist assoziativ, die identische Abbildung $\mathrm{id}_V : V \to V$ ist das neutrale Element bezüglich der Komposition in $\mathrm{Aut}(V)$ und die Umkehrabbildung f^{-1} ist das inverse Element zu $f \in \mathrm{Aut}(V)$. Man nennt $\mathrm{Aut}(V)$ auch die *allgemeine lineare Gruppe* von V, die diesem Namen entsprechende Bezeichnung ist $\mathrm{Gl}(V)$.

10.17. SATZ. *Sei V ein Vektorraum über K. Dann ist $\mathrm{End}(V)$ zusammen mit Addition $+$, skalarer Multiplikation $\cdot$ und Komposition $\circ$ eine* Algebra über K. *Es gilt nämlich*

(1) $(\mathrm{End}(V), +, \cdot)$ *ist ein Vektorraum über K.*
(2) $(\mathrm{End}(V), +, \circ)$ *ist ein Ring mit Eins. Das neutrale Element der Komposition ist die identische Abbildung id_V, und die Gruppe der Einheiten ist $\mathrm{Aut}(V)$.*
(3) *Für alle $f, g \in \mathrm{End}(V)$ und $\alpha \in K$ ist $\alpha \cdot (f \circ g) = (\alpha f) \circ g = f \circ (\alpha g)$.*

Proof. Aussage (1) ist schon bewiesen. Wir überprüfen ein Distributivgesetz:

$$(f \circ (g + h))(v) = f((g + h)(v)) = f(g(v) + h(v)) = (f \circ g)(v) + (f \circ h)(v)$$

für alle $v \in V$, daher ist $f \circ (g+h) = (f \circ g) + (f \circ h)$. Die Überprüfung der restlichen Eigenschaften ist eine ähnlich einfache Trockenübung. □

Wir diskutieren jetzt die Komposition für die wichtigste Beispielklasse, nämlich für die durch Matrizen gegebenen linearen Abbildungen. Seien $f : K^n \to K^m$ und $g : K^m \to K^l$ lineare Abbildungen, gegeben durch die Matrizen $A \in \mathrm{Mat}(m \times n, K)$ und $B \in \mathrm{Mat}(l \times m, K)$. Für $x \in K^n$ und $y \in K^m$ ist dann $f(x) = Ax$ und $g(y) = By$, also $(g \circ f)(x) = B(Ax)$. Sei jetzt $y = Ax$. Die i-te Komponente von By ist $\sum_{j=1}^m \beta_{ij} y_j$. Die j-te Komponente von $y = Ax$ ist $\sum_{k=1}^n \alpha_{jk} x_k$. Insgesamt ist daher die i-te Komponente von $B(Ax)$ gegeben durch

$$(*) \qquad \sum_{j=1}^m \beta_{ij} \Big(\sum_{k=1}^n \alpha_{jk} x_k \Big) = \sum_{j=1}^m \sum_{k=1}^n \beta_{ij} \alpha_{jk} x_k = \sum_{k=1}^n \Big(\sum_{j=1}^m \beta_{ij} \alpha_{jk} \Big) x_k.$$

In den Klammern rechts steht der Eintrag γ_{ik} der Matrix C von $g \circ f$. Die Matrix C ist das *Produkt* von B und A,

$$(10.18) \qquad C := BA. \qquad \textbf{Achtung:} \text{ Die Reihenfolge spielt eine Rolle!}$$

Die Regel zur Berechnung des Produkts entnehmen wir $(*)$, sie ist analog zu der Regel für die Berechnung des Produkts einer Matrix mit einem Vektor: Zur Berechnung des Eintrags γ_{ik} des Produkts lege die k-te Spalte von A auf die i-te Zeile von B — beide haben Länge m —, multipliziere die sich entsprechenden Einträge und bilde die

Summe über diese Produkte! Bei der Berechnung des Eintrags γ_{ik} des Produkts BA spielen damit nur die i-te Zeile von B und die k-te Spalte von A eine Rolle.

Die identische Abbildung wird durch die Einheitsmatrix E_n repräsentiert: In der j-ten Spalte steht der j-te Einheitsvektor von K^n. Aus Symmetriegründen gilt die analoge Aussage auch für die Zeilen von E_n. (Skizziere E_n.)

Wir nennen eine Matrix $A \in \mathrm{Mat}(m \times n, K)$ *invertierbar*, wenn die zugehörige lineare Abbildung $K^n \to K^m$ bijektiv ist. Dazu muss $m = n$ sein. Falls die Matrix $B \in \mathrm{Mat}(n \times n, K)$ die Umkehrabbildung repräsentiert, so hat man $AB = BA = E_n$. Die Menge der invertierbaren $(n \times n)$-Matrizen heisst *allgemeine lineare Gruppe* und wird mit $\mathrm{Gl}(n)$ bezeichnet.

Mit unseren vorgehenden Bemerkungen zu $\mathrm{End}(V)$ und $\mathrm{Aut}(V)$ ist der folgende Satz klar.

10.19. SATZ. *Zusammen mit Addition, skalarer Multiplikation und Multiplikation von Matrizen ist* $\mathrm{Mat}(n \times n, K)$ *eine Algebra über* K. *Die Einheitsmatrix ist das neutrale Element der Matrizenmultiplikation,* $\mathrm{Gl}(n)$ *ist die Gruppe der Einheiten.*

11. Kern, Bild und Rang

Seien V und W Vektorräume über K und $f : V \to W$ linear. Dann nennen wir

$$\begin{aligned} \operatorname{Ker} f &:= f^{-1}(0) = \{x \in V \mid f(x) = 0\}, \\ \operatorname{Im} f &:= f(V) = \{y \in W \mid \text{es gibt } x \in V \text{ mit } f(x) = y\}, \\ \operatorname{Rang} f &:= \dim \operatorname{Im} f. \end{aligned} \tag{11.1}$$

den *Kern*, das *Bild* und den *Rang* von f. Eine erste wichtige Bemerkung zum Kern:

11.2. Lemma. *Sei $v \in V$ und $w = f(v) \in W$. Dann besteht das Urbild von w unter f genau aus den Vektoren $v+u$ mit $u \in \operatorname{Ker} f$. Insbesondere folgt: f ist injektiv genau dann, wenn $\operatorname{Ker} f = \{0\}$ ist.*

Beweis. Sei nämlich $u \in V$. Dann ist $f(v+u) = f(v) + f(u)$, also ist $f(v+u) = w$ genau dann, wenn $u \in \operatorname{Ker} f$. □

11.3. Satz. *Falls U ein Unterraum von V ist, so ist $f(U)$ ein Unterraum von W. Falls umgekehrt U ein Unterraum von W ist, so ist $f^{-1}(U)$ ein Unterraum von V. Insbesondere gilt: Kern und Bild von f sind Unterräume von V bzw. W.*

Beweis. Sei U ein Unterraum von V. Seien $y, y' \in f(U)$ und $\alpha \in K$. Wähle $x, x' \in U$ mit $f(x) = y, f(x') = y'$. Dann sind $y + y' = f(x) + f(x') = f(x + x') \in f(U)$ und $\alpha y = \alpha f(x) = f(\alpha x) \in f(U)$, denn U ist ein Unterraum von V.

Sei nun U ein Unterraum von W. Seien $x, x' \in f^{-1}(U)$ und $\alpha \in K$. Dann sind $f(x + x') = f(x) + f(x') \in U$ und $f(\alpha x) = \alpha f(x) \in U$, denn U ist ein Unterraum von W. Daher sind $x + x'$ und αx in $f^{-1}(U)$. □

11.4. Satz (Rangformel für lineare Abbildungen). *Sei $f : V \to W$ linear. Falls dann $\dim V < \infty$ ist, so ist*

$$\operatorname{Rang} f + \dim \operatorname{Ker} f = \dim V.$$

Beweis. Sei $(v_1, \dots, v_n)$ eine Basis von V. Sei $w \in \operatorname{Im} f$, also $w = f(v)$ für ein geeignetes $v \in V$. Dann kann v linear aus den v_i kombiniert werden, $v = \sum \alpha_i v_i$, somit

$$w = f(v) = f\big(\sum \alpha_i v_i\big) = \sum \alpha_i f(v_i).$$

Daher ist $(f(v_1), \dots, f(v_n))$ ein Erzeugendensystem von $\operatorname{Im} f$. Insbesondere folgt $\operatorname{Rang} f = \dim \operatorname{Im} f < \infty$.

Sei nun $(w_1, \dots, w_r)$ eine Basis von $\operatorname{Im} f$, $r = \operatorname{Rang} f$. Seien $u_1, \dots, u_r \in V$ Vektoren mit $f(u_i) = w_i$, $1 \le i \le r$. Dann sind die u_i linear unabhängig, also Basis ihrer linearen Hülle $U := L(u_1, \dots, u_r)$.

Sei nun $v \in V$ und $w = f(v)$. Dann kann w linear aus den w_i kombiniert werden, $w = \sum \beta_i w_i$, also

$$w = \sum \beta_i w_i = \sum \beta_i f(u_i) = f\big(\sum \beta_i u_i\big) = f(u)$$

mit $u = \sum \beta_i u_i \in U$. Also ist

$$f(v - u) = f(v) - f(u) = w - w = 0,$$

daher ist $v - u \in \operatorname{Ker} f$, damit $U + \operatorname{Ker} f = V$.

Sei nun $v \in U \cap \operatorname{Ker} f$, $v = \sum \gamma_i u_i$. Dann ist

$$0 = f(v) = f\big(\sum \gamma_i u_i\big) = \sum \gamma_i f(u_i) = \sum \gamma_i w_i.$$

Nun ist $(w_1, \dots, w_r)$ Basis von $\operatorname{Im} f$, ist also insbesondere linear unabhängig. Damit folgt $\gamma_i = 0$, $1 \le i \le r$, also $v = 0$. Mit der Dimensionsformel für Unterräume folgt

$$\dim U + \dim \operatorname{Ker} f = \dim V,$$

also die Behauptung des Satzes, denn $\dim U = r = \operatorname{Rang} f$. □

Sei $A \in \operatorname{Mat}(m \times n, K)$. Wie gehabt interpretieren wir A als lineare Abbildung $K^n \to K^m$ und nennen

$$\begin{aligned} \operatorname{Ker} A &:= \{x \in K^n \mid Ax = 0\}, \\ \operatorname{Im} A &= \{y \in K^m \mid \text{es gibt } x \in K^n \text{ mit } Ax = y\}, \\ \operatorname{Rang} A &:= \dim \operatorname{Im} A \end{aligned} \tag{11.5}$$

den *Kern*, das *Bild* und den *Rang* von A. Der Kern von A besteht also genau aus den Lösungen des *homogenen* linearen Gleichungssystems $Ax = 0$. Das Bild von A besteht genau aus den $y \in K^m$, für die das lineare Gleichungssystem $Ax = y$ lösbar ist.

11.6. KOROLLAR. *Sei $x_0 \in K^n$ und $y = Ax_0 \in K^m$. Die Lösungsmenge des linearen Gleichungssystems $Ax = y$ besteht dann genau aus den Vektoren $x_0 + u$, wobei $u \in K^n$ eine Lösung des homogenen linearen Gleichungssystems $Ax = 0$ ist.* □

11.7. KOROLLAR (Rangformel für Matrizen). *Sei $A \in \operatorname{Mat}(m \times n, K)$. Dann ist*

$$\operatorname{Rang} A + \dim \operatorname{Ker} A = n. \quad \square$$

Nun ist $\operatorname{Im} A$ ein Unterraum von K^m, also ist $\operatorname{Rang} A \le m$. Ferner ist $\dim \operatorname{Ker} A \ge 0$, also ist $\operatorname{Rang} A \le n$. Analoge Argumente gelten für lineare Abbildungen zwischen (endlichdimensionalen) Vektorräumen über K.

11.8. SATZ. *Seien V und W endlichdimensionale Vektorräume über K. Für eine lineare Abbildung $f : V \to W$ ist dann* $\operatorname{Rang} f \le \min(\dim V, \dim W)$. *Für eine Matrix $A \in \operatorname{Mat}(m \times n, K)$ gilt entsprechend* $\operatorname{Rang} A \le \min(m, n)$. □

11.9. KOROLLAR. *Seien U, V und W endlichdimensionale Vektorräume über K. Seien $f : U \to V$ und $g : V \to W$ lineare Abbildungen. Dann ist*

(1) $\operatorname{Rang}(g \circ f) \le \min(\operatorname{Rang} f, \operatorname{Rang} g)$.
(2) *Falls f surjektiv ist, so ist* $\operatorname{Rang}(g \circ f) = \operatorname{Rang} g$.
(3) *Falls g injektiv ist, so ist* $\operatorname{Rang}(g \circ f) = \operatorname{Rang} f$.

Entsprechend gilt für $A \in \operatorname{Mat}(m \times n, K)$ und $B \in \operatorname{Mat}(l \times m, K)$

(1) $\operatorname{Rang}(BA) \le \min(\operatorname{Rang} A, \operatorname{Rang} B)$.
(2) *Falls A surjektiv ist, so ist* $\operatorname{Rang}(BA) = \operatorname{Rang} B$.
(3) *Falls B injektiv ist, so ist* $\operatorname{Rang}(BA) = \operatorname{Rang} A$.

Beweis. Die erste Ungleichung folgt mit $g \circ f = (g|\operatorname{Im} f) \circ f$. Falls f surjektiv ist, so ist $\operatorname{Im}(g \circ f) = \operatorname{Im} g$. Falls g injektiv ist, so ist $\operatorname{Rang}(g \circ f) = \dim(g(f(V)) = \dim f(V) = \operatorname{Rang} f$. □

Weiter oben haben wir die Spalten $a_1, \ldots, a_n$ einer Matrix $A \in \operatorname{Mat}(m \times n, K)$ als Vektoren des K^m aufgefasst, der j-te Spaltenvektor ist $a_j = (\alpha_{1j}, \ldots, \alpha_{mj})$. Analog fassen wir die Zeilen von A als Vektoren des K^n auf, der i-te Zeilenvektor von A ist $(\alpha_{i1}, \ldots, \alpha_{in})$.

11.10. DEFINITION. Das maximale r, so dass das n-Tupel $(a_1, \ldots, a_n)$ in K^m ein linear unabhängiges Teiltupel $(a_{i_1}, \ldots, a_{i_r})$ enthält, nennen wir den *Spaltenrang* von A. Entsprechend definieren wir den *Zeilenrang* von A über das m-Tupel der Zeilenvektoren von A.

Die Spaltenvektoren $a_1, \ldots, a_n$ erzeugen das Bild von A, denn für $x \in K^n$ ist

$$Ax = x_1 a_1 + \ldots + x_n a_n.$$

Daher ist ein maximales linear unabhängiges Teiltupel der Spaltenvektoren eine Basis von $\operatorname{Im} A$. Damit folgt Spaltenrang $A = \dim \operatorname{Im} A = \operatorname{Rang} A$.

11.11. DEFINITION. Sei $A \in \operatorname{Mat}(m \times n, K)$. Eine *elementare Zeilenumformung* von A ist eine der folgenden drei Manipulationen:

(1) Vertauschung zweier Zeilen von A.
(2) Multiplikation einer Zeile von A mit einem Skalar $\lambda \in K, \lambda \neq 0$.
(3) Addition eines skalaren Vielfachen einer Zeile zu einer *anderen* Zeile.

Analog sind elementare *Spaltenumformungen* definiert.

Als Beispiel manipulieren wir die Telefonmatrix:

$$\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix} \quad \to \quad \begin{pmatrix} 4 & 8 & 12 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix} \quad \to \quad \begin{pmatrix} 4 & 8 & 12 \\ 0 & -3 & -6 \\ 7 & 8 & 9 \end{pmatrix} \quad \to$$

$$\to \quad \begin{pmatrix} 7 & 14 & 21 \\ 0 & -3 & -6 \\ 7 & 8 & 9 \end{pmatrix} \quad \to \quad \begin{pmatrix} 7 & 14 & 21 \\ 0 & -3 & -6 \\ 0 & -6 & -12 \end{pmatrix} \quad \to \quad \begin{pmatrix} 7 & 14 & 21 \\ 0 & -3 & -6 \\ 0 & 0 & 0 \end{pmatrix}.$$

Sei $U \subset K^n$ die lineare Hülle der Zeilenvektoren von A. Nach Definition ist Zeilenrang $A = \dim U$. Offenbar ändert sich U nicht unter elementaren Zeilenumformungen. Der Zeilenrang ändert sich daher bei elementaren Zeilenumformungen ebenfalls nicht. Der Zeilenrang der Telefonmatrix ist daher 2, denn der Zeilenrang der letzten Matrix in der Folge oben ist offenbar $= 2$.

Die elementaren Zeilenumformungen erhält man durch Multiplikation von A mit geeigneten $(m \times m)$-Matrizen von links:

Typ 1: Seien $i, j \in \{1, \ldots, m\}$, $i \neq j$. Sei $T \in \operatorname{Mat}(m \times m, K)$ die Matrix mit k-tem

Zeilenvektor

$$\tau_k = \begin{cases} e_k & \text{für } k \neq i, j. \\ e_i & \text{für } k = j, \\ e_j & \text{für } k = i, \end{cases} \tag{11.12}$$

wobei die e_k die Einheitsvektoren in K^n bezeichnen, $1 \leq k \leq n$. Dann entsteht TA aus A durch Vertauschung von i-ter und j-ter Zeile.
Typ 2: Seien $i \in \{1, \ldots, m\}$ und $\lambda \in K$. Sei S die Matrix mit k-tem Zeilenvektor

$$\sigma_k = \begin{cases} e_k & \text{für } k \neq i. \\ \lambda e_i & \text{für } k = i. \end{cases} \tag{11.13}$$

Dann entsteht SA aus A durch Multiplikation der i-ten Zeile mit λ.
Typ 3: Seien $i, j \in \{1, \ldots, m\}$, $i \neq j$, und $\lambda \in K$. Sei R die Matrix mit k-tem Zeilenvektor

$$\rho_k = \begin{cases} e_k & \text{für } k \neq i, \\ e_i + \lambda e_j & \text{für } k = i. \end{cases} \tag{11.14}$$

Dann erhält man RA aus A, indem man das λ-fache der j-ten Zeile zur i-ten Zeile addiert.
11.15. ÜBUNG. Skizziere die Matrizen R, S und T.
11.16. LEMMA. *Die Matrizen R, S (für $\lambda \neq 0$) und T sind invertierbar, ihre Inversen sind jeweils Matrizen desselben Typs.*

Beweis. Nach zweimaligem Vertauschen der i-ten und j-ten Zeile einer Matrix $A \in \mathrm{Mat}(m \times n, K)$ erhält man wieder die ursprüngliche Matrix A. Daher ist $T^{-1} = T$. Die inverse Matrix zu S erhält man, indem man den Eintrag $\lambda (\neq 0)$ durch $1/\lambda$ ersetzt. Die Matrix, die das λ-fache der j-ten Zeile von A von der i-ten Zeile abzieht, ist die zu der Matrix R oben inverse Matrix. □

11.17. KOROLLAR. *Bei elementaren Zeilenumformungen ändert sich der Rang von A nicht.*
11.18. DEFINITION. Sei $A \in \mathrm{Mat}(m \times n, K)$. Wir sagen, dass $A \in \mathrm{Mat}(m \times n, K)$ in *Zeilenstufenform* ist, wenn es eine Zahl $r \in \{0, \ldots, m\}$ gibt, so dass folgendes gilt:

(1) In den ersten r Zeilen von A ist jeweils mindestens ein Eintrag $\neq 0$. In den letzten $m - r$ Zeilen sind alle Einträge $= 0$.
(2) Für $1 \leq i \leq r$ sei $j(i) = \min\{j | a_{ij} \neq 0\}$. Dann ist $j(1) < j(2) < \ldots < j(r)$.

11.19. LEMMA. *Falls $A \in \mathrm{Mat}(m \times n, K)$ in Zeilenstufenform ist mit $r \in \{0, \ldots, m\}$ wie in der Definition, so ist* Spaltenrang A = Zeilenrang A = $\mathrm{Rang}\, A = r$.

Beweis. Seien $\alpha_1, \ldots, \alpha_r$ die ersten r Zeilen von A. Dann ist $\alpha_r \neq 0$, also das 1-Tupel (α_r) linear unabhängig. Sei nun $1 \leq i \leq r - 1$ und das $(r - i)$-Tupel $(\alpha_{i+1}, \ldots, \alpha_r)$ linear unabhängig. Nun ist α_i sicher nicht in der linearen Hülle dieses Tupels, denn $\alpha_{kj(i)} = 0$ für alle $k > i$, aber $a_{ij(i)} \neq 0$. Damit folgt, dass auch das Tupel $(\alpha_i, \ldots, \alpha_r)$

linear unabhängig ist. Per Induktion folgt, dass $(\alpha_1, \ldots, \alpha_r)$ linear unabhängig ist. Daher ist Zeilenrang $A = r$.

Offenbar ist der Rang von $A \leq r$. Andererseits sieht man sofort, dass die Gleichung $Ax = y$ genau dann lösbar ist, wenn $y_{r+1} = \ldots = y_m = 0$ ist. Damit folgt $\operatorname{Rang} A = r$. Mit Spaltenrang $A = \operatorname{Rang} A$ folgt das Lemma. □

11.20. SATZ. *Jede Matrix lässt sich mit elementaren Zeilenumformungen in Zeilenstufenform bringen.*

Beweis. Das folgende Schema führt zu einem Algorithmus (Übung), der eine gegebene Matrix $A \in \operatorname{Mat}(m \times n, K)$ in Zeilenstufenform bringt:

(1) Falls $A = 0$, so ist A schon in Zeilenstufenform.

(2) Falls $A \neq 0$, wähle ein $i \in \{1, \ldots, m\}$, so dass

$$j(i) = \min\{j | a_{ij} \neq 0\}$$

minimal ist. Falls $i > 1$, vertausche erste Zeile und i-te Zeile von A. Nach der Vertauschung hat (die neue Matrix) A die Eigenschaft, dass alle Spalten vor der $j(1)$-ten Spalte $= 0$ sind.

(3) Für eine Matrix A mit der in (2) zuletzt ausgesprochenen Eigenschaft, multipliziere die erste Zeile mit $1/a_{1j(1)}$. Damit wird $a_{1j(1)} = 1$.

(4) Sei A eine Matrix mit der in (2) zuletzt ausgesprochenen Eigenschaft und $a_{1j(1)} = 1$. Für alle $k \in \{2, \ldots, m\}$, subtrahiere das $a_{kj(1)}$-fache der ersten Zeile von der k-ten Zeile. Nach diesen Manipulationen hat (die neue Matrix) A die Eigenschaft, dass alle Einträge in der $j(1)$-ten Spalte ab der zweiten Zeile einschliesslich $= 0$ sind. Das heisst, A hat nun die Gestalt

$$\left(\begin{array}{cccc|c} 0 & \cdots & 0 & 1 & *\cdots* \\ \hline 0 & \cdots & 0 & 0 & \\ \vdots & & \vdots & \vdots & \tilde{A} \\ 0 & \cdots & 0 & 0 & \end{array}\right),$$

wobei die ersten $j(1) - 1$ Spalten $= 0$ sind, die $j(1)$-te Spalte ist der erste Einheitsvektor in K^m. Die Sternchen deuten an, dass uns diese Einträge nicht interessieren. Addition von Vielfachen einer der Zeilen 2 bis m ändern den $j(1)$-ten Eintrag der ersten Zeile nicht mehr. Die Matrix $\tilde{A}$ hat $m - 1$ Zeilen und $n - j(1)$ Spalten.

(5) Wende dasselbe Verfahren auf $\tilde{A}$ an.

Die Matrix, die wir mit diesem Verfahren schlussendlich erhalten, ist nicht nur in Zeilenstufenform, sondern hat auch jeweils $a_{ij(i)} = 1$ für die Zeilen, die nicht identisch $= 0$ sind. □

11.21. KOROLLAR. *Für alle $A \in \operatorname{Mat}(m \times n, K)$ gilt*

$$\text{Spaltenrang } A = \text{Zeilenrang } A = \operatorname{Rang} A.$$

Proof. Bei elementaren Zeilenumformungen ändern sich die drei Ränge nicht. Jede Matrix kann aber durch solche Umformungen in Zeilenstufenform gebracht werden. Für diese gilt die behauptete Gleichheit. □

Wir können die Zeilenstufenform noch normieren: Wir verlangen

$$a_{ij(i)} = 1 \quad \text{und} \quad a_{kj(i)} = 0 \text{ für } k \neq i,\ 1 \leq i, k \leq m. \tag{11.22}$$

Die ersten $j(3)$ Spalten der Matrix haben dann folgende Gestalt

$$\begin{pmatrix} 0\cdots 0 & 1 & *\cdots * & 0 & *\cdots * & 0 \\ & & & 1 & *\cdots * & 0 \\ & & & & & 1 \\ & & & & & 0 \\ & & & & & \vdots \\ & & & & & 0 \end{pmatrix},$$

in diesem Sinne geht es mit den restlichen Spalten weiter. Die südwestlichen Einträge sind alle $= 0$, die mit einem Sternchen bezeichneten Einträge lassen sich nicht mehr verbessern.

Offenbar gelangt man mit dem im Beweis von Satz 11.20 entwickelten Verfahren zu einer normierten Zeilenstufenform.

Eine Matrix $A \in \mathrm{Mat}(n \times n, K)$ ist invertierbar genau dann, wenn $\mathrm{Rang}\, A = n$ ist. Das Verfahren aus Satz 11.20 führt daher bei A zur Einheitsmatrix E_n, der einzigen Matrix in $\mathrm{Mat}(n \times n, K)$ mit Rang n, die in normierter Zeilenstufenform ist.

11.23. KOROLLAR. *Jede invertierbare Matrix $A \in \mathrm{Mat}(n \times n, K)$ ist ein Produkt von Matrizen der Form* (11.12)–(11.14).

Beweis. Sei $A \in \mathrm{Mat}(n \times n, K)$ invertierbar. Dann ist auch A^{-1} invertierbar, das Verfahren aus Satz 11.20 führt damit auch bei A^{-1} zur Einheitsmatrix. Jede dieser Umformungen wird durch Multiplikation von links mit einer geeigneten Matrix aus (11.12)–(11.14) bewirkt. Also ist

$$E_n = B_k \cdot (B_{k-1} \cdot \ldots \cdot (B_1 \cdot A^{-1}) \ldots) = (B_k \cdot \ldots \cdot B_1) \cdot A^{-1}$$

mit geeigneten Matrizen $B_1, \ldots, B_k$ der Typen (11.12)–(11.14). Damit folgt $A = B_k \cdot \ldots \cdot B_1$. □

Bei genauerer Betrachtung sieht man auch, dass man mit dem Verfahren aus Satz 11.20 eine obere Schranke für die Anzahl der Matrizen der Typen (11.12)–(11.14) erhält, die für die Darstellung von A als Produkt solcher Matrizen erforderlich ist.

12. Lineare Gleichungssysteme

Seien $A \in \mathrm{Mat}(m \times n, K)$ und $b \in K^m$. Wir betrachten das lineare Gleichungssystem $Ax = b$. Mit $S(A, b)$ bezeichnen wir die Lösungsmenge,

$$S(A, b) = \{x \in K^n \mid Ax = b\}. \tag{12.1}$$

Wir wissen schon: $S(A, 0) = \mathrm{Ker}\, A$ ist ein Unterraum in K^n und

$$S(A, b) = x_0 + S(A, 0) = \{x_0 + x \mid x \in S(A, 0)\}, \tag{12.2}$$

falls x_0 irgendeine Lösung von $Ax = b$ ist.

Die *erweiterte Matrix* zu dem linearen Gleichungssystem $Ax = b$ ist die Matrix $(A, b) \in \mathrm{Mat}(m \times (n+1), K)$, die aus A durch Hinzufügen von b als letzter Spalte entsteht,

$$(A, b) = \begin{pmatrix} \alpha_{11} & \alpha_{12} & \dots & \alpha_{1n} & \beta_1 \\ \alpha_{21} & \alpha_{22} & \dots & \alpha_{2n} & \beta_2 \\ \vdots & \vdots & & \vdots & \vdots \\ \alpha_{m1} & \alpha_{m2} & \dots & \alpha_{mn} & \beta_m \end{pmatrix}. \tag{12.3}$$

Seien $a_1, \cdots, a_n$ die Spaltenvektoren von A. Dann ist $x \in S(A, b)$ genau dann, wenn

$$x_1 a_1 + \cdots + x_n a_n = b.$$

Daher ist $Ax = b$ lösbar genau dann, wenn b in der linearen Hülle von $(a_1, \cdots, a_n)$ ist.

12.4. Satz. $S(A, b) \neq \emptyset \iff \mathrm{Rang}\, A = \mathrm{Rang}(A, b)$. □

12.5. Satz. *Bei elementaren Zeilenumformungen der Matrix (A, b) ändert sich die Lösungsmenge des zugehörigen linearen Gleichungssystems nicht.*

Beweis. Betrachte das Gleichungssystem $Ax = b$. Sei $B \in \mathrm{Mat}(m \times m, K)$ eine der Matrizen aus (11.12)–(11.14), die durch Multiplikation von links eine elementare Zeilenumformung bewirken. Dann ist B invertierbar, also

$$Ax = b \iff (BA)x = Bb.$$

Nun ersetzt die elementare Zeilenumformung gerade die Matrix (A, b) durch die neue Matrix $B \cdot (A, b) = (BA, Bb)$. □

Den ganzen Beweis noch einmal "zu Fuß": Wir schreiben das Gleichungssystem aus,

$$\begin{aligned} \alpha_{11}x_1 + \alpha_{12}x_2 + \ldots + \alpha_{1n}x_n &= \beta_1, \\ \alpha_{21}x_1 + \alpha_{22}x_2 + \ldots + \alpha_{2n}x_n &= \beta_2, \\ \vdots \qquad\quad & \vdots \qquad\qquad \vdots \quad \vdots \\ \alpha_{m1}x_1 + \alpha_{m2}x_2 + \ldots + \alpha_{mn}x_n &= \beta_m. \end{aligned} \tag{12.6}$$

Dann können wir offensichtlich zwei dieser Gleichungen vertauschen, ohne die Lösungsmenge zu verändern. Dabei tauschen wir ja nicht nur die Zeilen von A aus, sondern die

gesamte Gleichung, also auch die Einträge von b. Analog diskutiert man elementare Zeilenumformungen der zwei anderen Typen.

Nun wissen wir: (A, b) läßt sich durch elementare Zeilenumformungen in Zeilenstufenform bringen. Diese Matrix ist von der Form $(A', b') \in \mathrm{Mat}(m \times (n+1), K)$. Der Punkt ist: Erstens ist $S(A, b) = S(A', b')$, zweitens ist $S(A', b')$ einfach zu bestimmen. Zur Illustration betrachten wir als Beispiel das Gleichungssystem

$$\begin{aligned} x_1 + 2x_2 + 3x_3 &= 1, \\ 4x_1 + 5x_2 + 6x_3 &= 2, \\ x_1 + x_2 &= 0. \end{aligned}$$

Matrix A und Vektor b sind

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 1 & 1 & 0 \end{pmatrix}, \quad b = \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix}.$$

Wir erlauben uns, mehrere Transformationen der erweiterten Matrix in einem Schritt durchzuführen, der Übersichtlichkeit halber ist die letzte Spalte durch einen vertiklaen Strich vom ersten Teil abgetrennt:

$$\left(\begin{array}{ccc|c} 1 & 2 & 3 & 1 \\ 4 & 5 & 6 & 2 \\ 1 & 1 & 0 & 0 \end{array}\right) \rightarrow \left(\begin{array}{ccc|c} 1 & 1 & 0 & 0 \\ 1 & 2 & 3 & 1 \\ 4 & 5 & 6 & 2 \end{array}\right) \rightarrow \left(\begin{array}{ccc|c} 1 & 1 & 0 & 0 \\ 0 & 1 & 3 & 1 \\ 0 & 1 & 6 & 2 \end{array}\right) \rightarrow$$

$$\rightarrow \left(\begin{array}{ccc|c} 1 & 1 & 0 & 0 \\ 0 & 1 & 3 & 1 \\ 0 & 0 & 3 & 1 \end{array}\right) \rightarrow \left(\begin{array}{ccc|c} 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1/3 \end{array}\right) \rightarrow \left(\begin{array}{ccc|c} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1/3 \end{array}\right),$$

es gibt also genau eine Lösung, $x_0 = (0, 0, 1/3)$.

Eine Matrix $A \in \mathrm{Mat}(n \times n, K)$ ist genau dann invertierbar, wenn es eine Matrix $A' \in \mathrm{Mat}(n \times n, K)$ gibt mit $AA' = E_n$, wobei $E_n \in \mathrm{Mat}(n \times n, K)$ die Einheitsmatrix bezeichnet . Die j-te Spalte $x = a'_j$ von A' löst damit das Gleichungssystem $Ax = e_j$, wobei $e_j \in K^n$ den j-ten Einheitsvektor bezeichnet.

Wir wissen auch, dass A genau dann invertierbar ist, wenn Rang $A = n$ ist, also wenn die normierte Zeilenstufenform von A die Einheitsmatrix E_n ist. Damit haben wir auch schon das Verfahren, A^{-1} zu bestimmen. Erweitere A um die Einheitsmatrix $E = E_n$ und forme die erweiterte Matrix $(A, E) \in \mathrm{Mat}(n \times 2n, K)$ so um, dass die linke Hälfte zur Einheitsmatrix E wird. Die zweite Hälfte ist dann die zu A inverse Matrix.

Achtung: Man erreicht das Ziel genau dann, wenn A invertierbar ist!

In unserem Beispiel oben funktioniert das Verfahren wie folgt:

$$\left(\begin{array}{ccc|ccc} 1 & 2 & 3 & 1 & 0 & 0 \\ 4 & 5 & 6 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 & 0 & 1 \end{array}\right) \quad \to \quad \left(\begin{array}{ccc|ccc} 1 & 1 & 0 & 0 & 0 & 1 \\ 1 & 2 & 3 & 1 & 0 & 0 \\ 4 & 5 & 6 & 0 & 1 & 0 \end{array}\right) \quad \to$$

$$\left(\begin{array}{ccc|ccc} 1 & 1 & 0 & 0 & 0 & 1 \\ 0 & 1 & 3 & 1 & 0 & -1 \\ 0 & 1 & 6 & 0 & 1 & -4 \end{array}\right) \quad \to \quad \left(\begin{array}{ccc|ccc} 1 & 1 & 0 & 0 & 0 & 1 \\ 0 & 1 & 3 & 1 & 0 & -1 \\ 0 & 0 & 3 & -1 & 1 & -3 \end{array}\right) \quad \to$$

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

Damit ist

$$A^{-1} = \begin{pmatrix} -2 & 1 & -1 \\ 2 & -1 & 2 \\ -1/3 & 1/3 & -1 \end{pmatrix}.$$

Eine Matrix $A \in \mathrm{Mat}(m \times n, K)$, aufgefasst als lineare Abbildung $K^n \to K^m$, ist genau dann

(1) injektiv, wenn $\mathrm{Rang}\, A = n$ ist bzw. eine Matrix $B \in \mathrm{Mat}(n \times m, K)$ mit $BA = E_n$ existiert.
(2) surjektiv, wenn $\mathrm{Rang}\, A = m$ ist bzw. eine Matrix $B \in \mathrm{Mat}(n \times m, K)$ mit $AB = E_m$ existiert.

Die Matrix B heisst dann eine zu A links- bzw. rechtsinverse Matrix. Diese sind im Allgemeinen nicht eindeutig durch A bestimmt.
Übung: Mit Verfahren wie dem zur Berechnung der inversen Matrix kann man links- und rechtsinverse Matrizen zu A bestimmen.

13. Basen und Matrizen

Sei V ein endlichdimensionaler Vektorraum über K und $B = (v_1, \dots, v_n)$ eine Basis von V. Die Wahl von B bestimmt eine lineare Abbildung

$$\varphi_B : K^n \to V, \quad \varphi_B(e_j) = v_j \text{ für } 1 \le j \le n, \tag{13.1}$$

wobei $(e_1, \dots, e_n)$ die Standardbasis des K^n bezeichnet. Nach Lemma 10.7 ist φ_B ein Isomorphismus.

Sei nun W ein weiterer Vektorraum über K und $C = (w_1, \dots, w_m)$ eine Basis von W. Sei $f : V \to W$ eine lineare Abbildung. Mit den gewählten Basen B und C können wir dann eine Matrix

$$M_B^C(f) = (\alpha_{ij}) \in \mathrm{Mat}(m \times n, K) \tag{13.2}$$

bestimmen: Für $1 \le j \le n$ kann $f(v_j)$ genau auf eine Weise linear aus der Basis C kombiniert werden,

$$f(v_j) = \sum \alpha_{ij} w_i. \tag{13.3}$$

Hier sehen wir schon die Einträge α_{ij} der Matrix $M_B^C(f)$. Mit anderen Worten: In der j-ten Spalte der Matrix $M_B^C(f)$ stehen die Koeffizienten des Vektors $f(v_j) \in W$ bezüglich der Basis C.

13.4. Satz. *Für je zwei Basen B von V und C von W ist die Abbildung*

$$M_B^C : \mathrm{Hom}(V, W) \to \mathrm{Mat}(m \times n, K), \quad f \mapsto M_B^C(f),$$

ist ein Isomorphismus von Vektorräumen. Ferner ist $f \circ \varphi_B = \varphi_C \circ M_B^C(f)$.

Die letzte Behauptung des Satzes ist per definitionem dazu äquivalent, dass das Diagramm

$$\begin{array}{ccc} V & \xrightarrow{f} & W \\ {\scriptstyle \varphi_B}\uparrow & & \uparrow{\scriptstyle \varphi_C} \\ K^n & \xrightarrow{M_B^C(f)} & K^m \end{array} \tag{13.5}$$

kommutativ ist:
Ob wir in K^n mit φ starten und in V mit f weiterlaufen, oder ob wir mit $M_B^C(f)$ und φ_C über K^m nach W laufen, das Ergebnis ist dasselbe.

Beweis des Satzes. Für $f, f' \in \mathrm{Hom}(V, W)$ und $1 \le j \le n$ ist

$$\begin{aligned} (f + f')(e_j) &= f(e_j) + f'(e_j) \\ &= \sum \alpha_{ij} w_i + \sum \alpha'_{ij} w_i = \sum (\alpha_{ij} + \alpha'_{ij}) w_i, \end{aligned}$$

die rechte Seite entspricht genau der Addition in $\mathrm{Mat}(m \times n, K)$. Daher ist M_B^C additiv. Die Homogenität beweist man analog.

Sei $A = (\alpha_{ij}) \in \mathrm{Mat}(m \times n, K)$. Nach Lemma 10.7 gibt es genau eine lineare Abbildung $f : V \to W$ mit $f(v_j) := \sum \alpha_{ij} w_i$, und für diese ist $M_B^C(f) = A$. Daher ist M_B^C bijektiv und damit ein Isomorphismus von Vektorräumen.

Für $x = (x_1, \dots, x_n) \in K^n$ ist

$$f(\varphi_B(x)) = f(\sum_j x_j v_j) = \sum_j x_j f(v_j) = \sum_{ij} (\alpha_{ij} x_j) w_i = \varphi_C(M_B^C(f)(x)),$$

also ist $f \circ \varphi_B = \varphi_C \circ M_B^C(f)$ wie behauptet. □

13.6. SATZ. *Seien U, V, W Vektorräume über K mit jeweiligen Basen $B = (u_1, \cdots, u_n)$, $C = (v_1, \cdots, v_m)$ und $D = (w_1, \cdots, w_l)$. Seien $f : U \to V$ und $g : V \to W$ linear. Dann ist*

$$M_B^D(g \circ f) = M_C^D(g) \cdot M_B^C(f).$$

Beweis. Mit $M_B^C(f) = (\alpha_{jk})$ und $M_C^D(g) = (\beta_{ij})$ ist

$$f(u_k) = \sum \alpha_{jk} v_j, \quad g(v_j) = \sum \beta_{ij} w_i.$$

Damit ist

$$\begin{aligned}(g \circ f)(u_k) = g(f(u_k)) &= g\big(\sum_j \alpha_{jk} v_j\big) \\ &= \sum_j \alpha_{jk} g(v_j) = \sum_j \alpha_{jk} \big(\sum_i \beta_{ij} w_i\big) \\ &= \sum_j \sum_i \alpha_{jk} \beta_{ij} w_i = \sum_i \big(\sum_j \beta_{ij} \alpha_{jk}\big) w_i,\end{aligned}$$

der Koeffizient von w_i ist also genau der Eintrag in i-ter Zeile und k-ter Spalte von $M_C^D(g) \cdot M_B^C(f)$. □

Wie zum vorherigen Satz gibt es auch hier ein Diagramm,

$$\begin{array}{ccccc} U & \xrightarrow{f} & V & \xrightarrow{g} & W \\ {\scriptstyle\varphi_B}\uparrow & & {\scriptstyle\varphi_C}\uparrow & & {\scriptstyle\varphi_D}\uparrow \\ K^n & \xrightarrow{M_B^C(f)} & K^m & \xrightarrow{M_C^D(g)} & K^l. \end{array} \tag{13.7}$$

Der Satz besagt, dass dieses Diagramm kommutativ ist.

13.8. KOROLLAR. *Sei V ein Vektorraum über K und $B = (v_1, \cdots, v_n)$ eine Basis von V. Dann ist*

$$M_B^B : \mathrm{End}(V) \to \mathrm{Mat}(n \times n, K)$$

ein Isomorphismus von Algebren über K.

13.9. BEISPIELE. 1) $M_B^B(\mathrm{id}_V) = E_n$.

2) Sei $V = \{p \in K[x] \mid \mathrm{Grad}\, p \le 3\}$, $W = K$ als Vektorraum über sich selber und $f : V \to W$ die *Auswertung in* α, wobei $\alpha \in K$ ein fest gewählter Skalar ist:

$$p = a_0 + a_1 x + a_2 x^2 + a_3 x^3 \mapsto f(p) := a_0 + a_1 \alpha + a_2 \alpha^2 + a_3 \alpha^3.$$

Diese Abbildung ist linear. Das Tupel $B = (1, x, x^2, x^3)$ ist eine Basis von V, das 1-Tupel $C = (1)$ ist eine Basis von $W = K$. Wir berechnen $M_B^C(f)$ zu $(1, \alpha, \alpha^2, \alpha^3)$, denn $f(x^i) = \alpha^i = \alpha^i \cdot 1$. Eine andere Basis von V ist

$$B_\alpha = (1, x - \alpha, (x - \alpha)^2, (x - \alpha)^3).$$

Offenbar ist $M_{B_\alpha}^C(f) = (1, 0, 0, 0)$.

13.10. SATZ. *Seien V und W Vektorräume über K mit $\dim V = n$ und $\dim W = m$. Sei $f : V \to W$ linear mit $\operatorname{Rang}(f) = r$. Dann gibt es Basen B von V und C von W mit*

$$M_B^C(f) = \left(\begin{array}{c|c} E_r & 0 \\ \hline 0 & 0 \end{array}\right),$$

wobei E_r die $(r \times r)$-Einheitsmatrix ist und die drei Nullen die Nullmatrizen in den entsprechenden Größen bezeichnen.

Beweis. Wähle eine Basis $(w_1, \ldots, w_r)$ von $\operatorname{Im} f$ und ergänze sie zu einer Basis $C = (w_1, \ldots, w_m)$ von W. Für $1 \leq j \leq r$ wähle $v_j \in V$ mit $f(v_j) = w_j$. Dann ist das r-Tupel $(v_1, \ldots, v_r)$ linear unabhängig, denn $(w_1, \ldots, w_r)$ ist linear unabhängig. Sei $U = L(v_1, \ldots, v_r)$. Dann ist $U \cap \operatorname{Ker} f = \{0\}$ und $U + \operatorname{Ker} f = V$. Wähle eine Basis $(v_{r+1}, \ldots, v_n)$ von $\operatorname{Ker} f$. Dann ist $B = (v_1, \ldots, v_r, v_{r+1}, \ldots, v_n)$ eine Basis von V und $M_B^C(f)$ ist nach Wahl der Basen von der behaupteten Form. □

Die beiden grundlegenden Regeln für den Umgang mit den Matrizen $M_B^C(f)$ sind

(1) $M_B^B(\mathrm{id}_V) = E_n$,
(2) $M_B^D(g \circ f) = M_C^D(g) M_B^C(f)$.

Diese beiden Regeln liefern bei vielen Überlegungen genau die brauchbaren (und richtigen) Schlüsse.

Seien V und W Vektorräume über K mit $\dim V = n$ und $\dim W = m$. Sei $f : V \to W$ linear, seien B und B' Basen von V und C und C' Basen von W. Dann ist

$$M_{B'}^{C'}(f) = M_C^{C'}(\mathrm{id}_W)\, M_B^C(f)\, M_{B'}^B(\mathrm{id}_V).$$

Die Matrizen $M_{B'}^B(\mathrm{id}_V) \in \operatorname{Mat}(n \times n, K)$ und $M_C^{C'}(\mathrm{id}_W) \in \operatorname{Mat}(m \times m, K)$ beschreiben die Basiswechsel in V und W. Mit

$$M_{B'}^B(\mathrm{id}_V)\, M_B^{B'}(\mathrm{id}_V) = M_B^B(\mathrm{id}_V) = E_n$$

folgt, dass $M_B^{B'}(\mathrm{id}_V)$ invertierbar ist mit

$$\left(M_B^{B'}(\mathrm{id}_V)\right)^{-1} = M_{B'}^B(\mathrm{id}_V).$$

Die Umkehrung des zu Anfang des Abschnitts definierten Isomorphismus $\varphi_B : K^n \to V$ nennen wir *Koordinaten (bezüglich B)* und bezeichnen sie mit κ_B. Mit

dieser Notation folgt, dass die folgenden Diagramme kommutativ sind:

$$\begin{array}{ccc} V & \xrightarrow{f} & W \\ {\scriptstyle\kappa_B}\downarrow & & {\scriptstyle\kappa_C}\downarrow \\ K^n & \xrightarrow{M_B^C(f)} & K^m \end{array} \quad \text{bzw.} \quad \begin{array}{ccccc} U & \xrightarrow{f} & V & \xrightarrow{g} & W \\ {\scriptstyle\kappa_B}\downarrow & & {\scriptstyle\kappa_C}\downarrow & & {\scriptstyle\kappa_D}\downarrow \\ K^n & \xrightarrow{M_B^C(f)} & K^m & \xrightarrow{M_C^D(g)} & K^l. \end{array} \tag{13.11}$$

Insbesondere ist

$$\begin{array}{ccc} V & \xrightarrow{\mathrm{id}_V} & V \\ {\scriptstyle\kappa_B}\downarrow & & {\scriptstyle\kappa_{B'}}\downarrow \\ K^n & \xrightarrow{M_B^{B'}(\mathrm{id}_V)} & K^n, \end{array} \tag{13.12}$$

also ist der *Koordinatenwechsel* $\kappa_{B'} \circ \kappa_B^{-1} = M_B^{B'}(\mathrm{id}_V)$.

14. Der Dualraum

Sei V ein Vektorraum über K. Wir nennen

$$V^* = \mathrm{Hom}(V, K) \tag{14.1}$$

den *Dualraum* von V, die Elemente φ aus V^* nennen wir *Linearformen (auf V)*.

14.2. Beispiele (zu Linearformen). 1) Alte Bekannte: Sei $V = K^n$. Zu Skalaren $\alpha_1, \ldots, \alpha_n$ definiere $\varphi : K^n \to K$ durch

$$\varphi(x) := \alpha_1 x_1 + \cdots + \alpha_n x_n \in K.$$

Dies entspricht unserer früheren Interpretation von $(1 \times n)$-Matrizen mit Einträgen in K, hier die Matrix $(\alpha_1, \cdots, \alpha_n)$, als lineare Abbildungen von K^n nach K.

2) Sei $V = K[x]$, der Vektorraum der Polynome mit Koeffizienten in K. Sei $\alpha \in K$. Definiere $\varphi : K[x] \to K$ die *Auswertung in* α, durch

$$\varphi(a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n) := a_0 + a_1\alpha + a_2\alpha^2 + \cdots + a_n\alpha^n.$$

Nach Definition der Addition und skalaren Multiplikation in $V = K[x]$ ist φ linear, damit in V^*. Diese Abbildung haben wir schon in Beispiel 13.9.2 auf einem Unterraum von $K[x]$ betrachtet.

3) Sei I ein Intervall in $\mathbb{R}$ und $V = C(I, \mathbb{R})$ der $\mathbb{R}$-Vektorraum der stetigen Funktionen auf I mit Werten in $\mathbb{R}$. Sei $a \in I$ und $\varphi : C(I, \mathbb{R}) \to \mathbb{R}$, die *Auswertung in* a,

$$\varphi(f) := f(a).$$

Addition und skalare Multiplikation in $C(I, \mathbb{R})$ sind punktweise erklärt, daher ist φ linear und damit Element von V^*. Diese Linearform wird auch *Delta-Funktion (in a)* genannt (und dann etwa mit δ_a bezeichnet).

Sei $B = (v_1, \cdots, v_n)$ eine Basis von V. Nach Lemma 10.7 ist eine Linearform auf V durch ihre Werte auf B festgelegt und diese können beliebig vorgeschrieben werden. Wir defineren damit ein n-Tupel $B^* = (v_1^*, \ldots, v_n^*) \in V^*$ durch

$$v_i^*(v_j) = \begin{cases} 1 & \text{falls } i = j, \\ 0 & \text{falls } i \neq j. \end{cases} \tag{14.3}$$

Wir nennen B^* die zu B *duale Basis*. Der folgende Satz rechtfertigt diese Benennung.

14.4. Satz. *Die lineare Abbildung $\psi_B : V \to V^*$, definiert durch $\psi_B(v_j) = v_j^*$ für $1 \leq j \leq n$, ist ein Isomorphismus. Anders ausgedrückt: $B^* = (v_1^*, \cdots, v_n^*)$ ist eine Basis von V^*.*

Beweis. Wir wissen schon, dass $V^* = \mathrm{Hom}(V, K)$ die Dimension $n = n \cdot 1 = \dim V \cdot \dim K$ hat. Wir müssen also nur die Surjektivität von ψ_B zeigen. Sei dazu $\varphi \in V^*$. Setze $\alpha_i := \varphi(v_i)$ und $\varphi' := \sum \alpha_i v_i^* = \psi_B\big(\sum \alpha_i v_i\big)$. Dann ist

$$\varphi'(v_j) = \sum \alpha_i v_i^*(v_j) = \alpha_j v_j^*(v_j) = \alpha_j = \varphi(v_j).$$

Nun ist B eine Basis von V, aus Lemma 10.7 folgt damit $\varphi' = \varphi$. Daher ist ψ_B surjektiv. □

Der Beweis des Satzes liefert auch das Rezept dafür, wie $\varphi \in V^*$ linear aus B^* kombiniert wird,

$$\varphi = \sum \alpha_i v_i^* \quad \text{mit } \alpha_i = \varphi(v_i). \tag{14.5}$$

Sei nun $f : V \to W$ linear. Dann definieren wir eine Abbildung $f^* : W^* \to V^*$, *die zu f duale* oder *transponierte Abbildung*, durch

$$f^*(\varphi) := \varphi \circ f. \tag{14.6}$$

14.7. SATZ. *Die zu f duale Abbildung $f^* : W^* \to V^*$ ist linear. Ferner ist*

(1) $(\mathrm{id}_V)^* = \mathrm{id}_{V^*}$,
(2) $(g \circ f)^* = f^* \circ g^*$. □

Man vergleiche die beiden Regeln des vorstehenden Satzes mit den Regeln (1) und (2) auf Seite 49.

Für eine Matrix $A \in \mathrm{Mat}(m \times n, K)$ definieren wir die *transponierte Matrix* $A^t \in \mathrm{Mat}(n \times m, K)$ durch

$$\alpha_{ij}^t := \alpha_{ji}, \tag{14.8}$$

die zu A transponierte Matrix entsteht also durch Spiegelung von A entlang der Diagonalen $i = j$. Die folgenden Regeln sind offensichtlich,

$$E_n^t = E_n \quad \text{und} \quad (AB)^t = B^t A^t. \tag{14.9}$$

Man vergleiche diese Regeln mit den Regeln (1) und (2) auf Seite 49 und aus Satz 14.7.

14.10. SATZ. *Falls V und W endlichdimensional sind und $B = (v_1, \dots, v_n)$ bzw. $C = (w_1, \dots, w_m)$ Basen von V bzw. W sind, so ist*

$$M_{C^*}^{B^*}(f^*) = (M_B^C(f))^t.$$

Beweis. Für $1 \le i \le m$ ist

$$f^*(w_i^*)(v_j) = (w_i^* \circ f)(v_j) = w_i^*(f(v_j)) = w_i^*\big(\sum \alpha_{kj} w_k\big) = \alpha_{ij},$$

also ist $f^*(w_i^*) = \sum \alpha_{ij} v_j^*$. □

Sei $M \subset V$ eine Teilmenge. Dann nennen wir

$$M^0 = \{\varphi \in V^* \mid \varphi(v) = 0 \text{ für alle } v \in M\} \tag{14.11}$$

den *Annihilator* von M.

14.12. SATZ. *Für alle Teilmengen $M \subset V$ ist M^0 ein Unterraum von V^*. Ferner gilt:*

(1) $V^0 = \{0\}$ *und* $\{0\}^0 = V^*$.
(2) $M \subset N \subset V \Longrightarrow N^0 \subset M^0$. □

14.13. SATZ (Dimensionsformel). *Sei $\dim V < \infty$ und $U \subset V$ ein Unterraum. Dann ist*

$$\dim U + \dim U^0 = \dim V.$$

Genauer gilt: Falls $B = (v_1, \dots, v_n)$ eine Basis von V ist, so dass das Teiltupel $(v_1, \dots, v_k)$ eine Basis von U ist, so ist das Teiltupel $(v_{k+1}^*, \dots, v_n^*)$ der dualen Basis B^* von V^* eine Basis von U^0.

Beweis. Sei $\varphi \in V^*$. Dann kann φ linear aus B^* kombiniert werden, $\varphi = \sum \alpha_i v_i^*$. Hierbei ist $\alpha_i = \varphi(v_i)$, also $\varphi \in U^0$ genau dann, wenn $\alpha_i = 0$ ist für $1 \leq i \leq k$. □

14.14. SATZ. *Seien V und W endlichdimensional und $f : V \to W$ linear. Dann gilt*

$$\operatorname{Im} f^* = (\operatorname{Ker} f)^0 \quad \textit{und} \quad \operatorname{Ker} f^* = (\operatorname{Im} f)^0.$$

Beweis. Man vergleiche den Beweis mit dem Beweis von Satz 13.10. Sei $\operatorname{Rang} f = r$ und $(w_1, \ldots, w_r)$ eine Basis von $\operatorname{Im} f$. Ergänze diese zu einer Basis $C = (w_1, \ldots, w_m)$ von W. Für $1 \leq j \leq r$ sei $v_j \in V$ ein Vektor mit $f(v_j) = w_j$. Dann ist das r-Tupel $(v_1, \ldots, v_r)$ linear unabhängig. Sei $U = L(v_1, \ldots, v_r)$. Dann ist $U \cap \operatorname{Ker} f = \{0\}$ und $U + \operatorname{Ker} f = V$. Wähle eine Basis $(v_{r+1}, \ldots, v_n)$ von $\operatorname{Ker} f$. Dann ist $(v_1, \ldots, v_n)$ eine Basis von V. Wir berechnen

$$f^*(w_i^*)(v_j) = w_i^*(f(v_j)) = \begin{cases} 0 & \text{falls } j > r, \\ w_i^*(w_j) & \end{cases} = \begin{cases} 0 & \text{falls } j \leq r \text{ und } i \neq j, \\ 1 & \text{falls } j \leq r \text{ und } i = j. \end{cases}$$

Damit ist $f^*(w_i^*) = v_i^*$ für $i \leq r$ und $f^*(w_i^*) = 0$ für $i > r$. Insbesondere ist $\operatorname{Ker} f^* = L(w_{r+1}^*, \ldots, w_m^*)$ und $\operatorname{Im} f^* = L(v_1^*, \ldots, v_r^*)$. Die Behauptungen des Satzes folgen nun aus Satz 14.13. □

14.15. KOROLLAR. *Seien V und W endlichdimensional und $f : V \to W$ linear. Dann ist* $\operatorname{Rang} f^* = \operatorname{Rang} f$.

Beweis. Mit Satz 14.13 und der Rangformel für lineare Abbildungen folgt

$$\operatorname{Rang} f^* = \dim \operatorname{Im} f^* = \dim (\operatorname{Ker} f)^0 = \dim V - \dim \operatorname{Ker} f = \operatorname{Rang} f.$$

□

14.16. KOROLLAR. *Zeilenrang und Spaltenrang von Matrizen stimmen überein.*
Sie erinnern sich richtig, das haben wir schon einmal — aber auf anderem Wege — bewiesen. Der neue Beweis zeigt, dass dieses Ergebnis ein Spezialfall von Korollar 14.15 ist. Auf unserem Weg auf der Spirale nach oben haben wir eine Windung zurückgelegt.

Beweis von Korollar 14.16. Sei $A \in \operatorname{Mat}(m \times n, K)$. Fasse A als lineare Abbildung $K^n \to K^m$ auf, das heisst, A ist die Matrixdarstellung dieser linearen Abbildung bezüglich der Standardbasen von K^n und K^m. Nach Satz 14.10 ist die Matrix der dualen Abbildung bezüglich der entsprechenden dualen Basen durch A^t gegeben. Nach Definition ist nun der Zeilenrang von A gleich dem Spaltenrang, also dem Rang, von A^t. Nach Korollar 14.15 ist dieser aber gleich dem Rang, also dem Spaltenrang, von A. □

Sei nun wieder V ein Vektorraum über K. Wir betrachten die natürliche *Paarung* zwischen V und V^*, also die Abbildung

$$V^* \times V \to K, \ (\varphi, v) \mapsto (\varphi | v) := \varphi(v). \tag{14.17}$$

Die Notation $(\varphi|v)$ soll andeuten, dass die Partner φ und v gleichberechtigt sind und damit ihre gegenseitige Beziehung harmonisch ist.

Nach Definition von V^* und der Vektorraumstruktur auf V^* ist die Paarung 14.17 *bilinear*, das heisst, für alle φ in V^* ist die Zuordnung $V \ni v \mapsto (\varphi|v) \in K$ linear und für alle $v \in V$ ist die Zuordnung $V^* \ni \varphi \mapsto (\varphi|v) \in K$ linear. Falls ferner $f : V \to W$ eine lineare Abbildung ist, so ist

$$(\psi|f(v)) = (f^*(\psi)|v) \tag{14.18}$$

für alle $v \in V$ und $\psi \in W^*$. Die Paarung verträgt sich daher mit Abbildungen und den zugehörigen dualen Abbildungen.

Im nächsten Satz betrachten wir Linearformen auf V^*, also Elemente aus dem sogenannten *Bidualraum* $V^{**} := (V^*)^*$, dem Dualraum des Dualraums. Die Aussage des Satzes ist, dass wir uns, solange V von endlicher Dimension ist, nicht weiter um V^{**} scheren müssen, damit erst recht nicht um noch höhere Stufen der Dualisierung.

14.19. SATZ. *Falls V endliche Dimension hat, so ist die Paarung 14.17* perfekt*: Zu jeder Linearform $\Psi : V^* \to K$ gibt es genau einen Vektor $v \in V$ mit $\Psi(\varphi) = (\varphi|v)$ für alle $\varphi \in V^*$.*

Beweis. Die Abbildung $F : V \to V^{**}$, die $v \in V$ die Linearform $V^* \ni \varphi \mapsto (\varphi|v) \in K$ auf V^* zuordnet, ist nach Definition der Vektorraumstruktur auf V^* linear. Zu $v \in V$ mit $v \neq 0$ gibt es ein $\varphi \in V^*$ mit $(\varphi|v) = \varphi(v) \neq 0$. Also ist $\operatorname{Ker} F = \{0\}$, daher ist F injektiv [18]. Nun ist $\dim V^* = \dim V < \infty$. Also ist $\dim V^{**} = \dim V^* = \dim V$. Deshalb ist F auch surjektiv. □

Sei nun $M \subset V^*$ eine Teilmenge. Dann nennen wir

$$M^0 := \{v \in V \mid (\varphi|v) = 0 \text{ für alle } \varphi \in M\} \tag{14.20}$$

den *Annihilator* von M. Die den Regeln in Satz 14.12 entsprechenden Regeln gelten auch hier.

14.21. SATZ. *Für alle Teilmengen $M \subset V^*$ ist M^0 ein Unterraum von V. Ferner gilt:*

(1) $(V^*)^0 = \{0\}$ *und* $\{0\}^0 = V$.

(2) $M \subset N \subset V^* \Longrightarrow N^0 \subset M^0$. □

Für $M \subset V$ bzw. $N \subset V^*$ können wir jetzt jeweils zweimal den Annihilator bilden, $M^{00} := (M^0)^0 \subset V$ bzw. $N^{00} := (N^0)^0 \subset V^*$.

14.22. SATZ. *Sei V von endlicher Dimension und seien $M \subset V$ und $N \subset V^*$. Dann ist M^{00} bzw. N^{00} die lineare Hülle von M bzw. N.*

Streng genommen haben wir lineare Hüllen von Mengen nicht definiert. Es ist aber offensichtlilch, was damit gemeint ist. Der Beweis von Satz 14.22 bleibt dem geneigten Leser überlassen.

[18] Das stimmt auch, wenn V unendliche Dimension hat; F ist dann aber nicht mehr surjektiv.

15. ANHANG

Sei X eine Menge. Eine *Relation* auf X ist eine Teilmenge R von $X \times X$.

15.1. DEFINITION. Eine Relation R auf X heisst *Äquivalenzrelation*, wenn sie die folgenden drei Eigenschaften hat:

(1) Für alle $x \in X$ ist $(x, x) \in R$.
(2) $(x, y) \in R \Rightarrow (y, x) \in R$.
(3) $(x, y) \in R$ und $(y, z) \in R \Rightarrow (x, z) \in R$.

Falls R eine Äquivalenzrelation ist, dann sagen wir, dass Elemente $x, y \in X$ *äquivalent* sind bezüglich R, wenn $(x, y) \in R$ ist. Wir schreiben dann auch einfach $x \sim y$, ohne R explizit zu erwähnen und nennen $\{y \in X \mid x \sim y\}$ die *Äquivalenzklasse* von x bezüglich R.

Die erste Eigenschaft in Definition 15.1 heisst *Reflexivität*, die zweite *Symmetrie*, die dritte *Transitivität*.

Eine *Partition* von X ist eine Menge $\mathcal{P}$ von Teilmengen von X, die paarweise disjunkt sind, und deren Vereinigung $\cup_{Y \in \mathcal{P}} Y = X$ ist.

15.2. SATZ. *Sei R eine Äquivalenzrelation auf X. Zu $x \in X$ sei R_x die Äquivalenzklasse von x. Dann gilt:*

(1) *Für alle $x \in X$ ist $x \in R_x$.*
(2) $y \in R_x \Leftrightarrow x \in R_y$.
(3) $x \sim y \Leftrightarrow R_x = R_y \Leftrightarrow R_x \cap R_y \neq \emptyset$.

Insbesondere ist die Menge der Äquivalenzklassen $\{R_x \mid x \in X\}$ eine Partition der Menge X.

Beweis. (1) folgt sofort aus der Reflexivität von R, (2) aus der Symmetrie.

Zum Beweis von (3) führen wir einen Kreisschluss durch. Sei nun $x \sim y$ und $z \in R_y$. Dann ist $y \sim z$, also $x \sim z$ wegen der Transitivität von R. Damit folgt $z \in R_x$, also $R_y \subset R_x$. Aus der Symmetrie von R folgt, dass wir die Rollen von x und y vertauschen dürfen, also gilt auch $R_x \subset R_y$. Mithin folgt $R_x = R_y$ aus $x \sim y$.

Wegen $x \in R_x$ ist R_x nicht leer. Damit impliziert $R_x = R_y$, dass $R_x \cap R_y \neq \emptyset$ ist.

Sei nun der Durchschnitt von R_x mit R_y nicht leer und $z \in R_x \cap R_y$. Dann ist $x \sim z$ und $y \sim z$, also wegen der Symmetrie auch $z \sim y$. Aus der Transitivität von R folgt $x \sim y$. Damit sind wir bei unserem Kreisschluss wieder da angelangt, wo wir angefangen haben, und sind doch um einiges klüger. □

Anzumerken bleibt, dass ein Kreisschluss kein circulus vitiosus ist. Es ist auch nicht so, dass sich die Katze bei einem Kreisschluss in den Schwanz beisst.

15.3. DEFINITION. Eine Relation R auf X heisst eine *Ordnung* von X, und X zusammen mit R heisst dann eine *geordnete Menge*, wenn R die folgenden drei Eigenschaften hat:

(1) Für alle $x \in X$ ist $(x, x) \in R$.
(2) $(x, y) \in R$ und $(y, x) \in R \Rightarrow x = y$.
(3) $(x, y) \in R$ und $(y, z) \in R \Rightarrow (x, z) \in R$.

Falls R eine Ordnungsrelation ist, so schreiben wir auch einfach $x \leq y$ oder $y \geq x$, falls $(x, y) \in R$ ist. Wir schreiben $x < y$ oder $y > x$, falls $x \leq y$ und $x \neq y$.

Als Beispiele geordneter Mengen seien $\mathbb{N}, \mathbb{Z}, \mathbb{Q}$ und $\mathbb{R}$ mit der üblichen $\leq$-Beziehung genannt.

REFERENCES

[Br] E. Brieskorn: *Lineare Algebra und analytische Geometrie*. I,II. Noten zu einer Vorlesung mit historischen Anmerkungen von E. Scholz. Vieweg & Sohn, Braunschweig, 1983 & 1985.

[Fi] G. Fischer: *Lineare Algebra*. Eine Einführung für Studienanfänger. Vieweg & Sohn, Braunschweig, 1975.

[Ga] P. Gabriel: *Matrizen, Geometrie, Lineare Algebra*. Birkhäuser Verlag, Basel [u.a.], 1996.

[Jä] K. Jänich: *Lineare Algebra*. Ein Skriptum für das erste Semester. Springer-Verlag, Berlin [u.a.], 1979.

[Kö] M. Koecher: *Lineare Algebra und analytische Geometrie*. Springer-Verlag, Berlin [u.a.], 1983.

[RU] R. Remmert und P. Ullrich: *Elementare Zahlentheorie*. Birkhäuser Verlag, Basel [u.a.], 1987.

MATHEMATISCHES INSTITUT, UNIVERSITÄT BONN, BERINGSTRASSE 1, 53115 BONN,
E-mail address: ballmann@math.uni-bonn.de