

Lineare Algebra

Vorlesung

von

Heinz Spindler

Fachbereich Mathematik/Informatik

Universität Osnabrück

Wintersemester 2004/05

Inhaltsverzeichnis

1 Lösen von Gleichungen	4
2 Lineare Gleichungssysteme, erste Schritte	12
3 Mengen und Abbildungen	23
4 Vollständige Induktion	36
5 Äquivalenzrelationen	41
6 Körper	46
7 Vektorräume	60
8 Basis und Dimension	67
9 Lineare Abbildungen	79
10 Der Hom-Funktor und Dualräume	86
11 Lineare Abbildungen und Matrizen	98
12 Der Rang einer Matrix: Äquivalenz von Matrizen	111
13 Lineare Gleichungssysteme und affine Unterräume	125
14 Ringe, Algebren, Polynomringe, Matrizenalgebren	134
15 Determinanten	148
16 Eigenwerte, Eigenvektoren, charakteristisches Polynom	161
17 Euklidische Vektorräume	174
18 Orthogonale Abbildungen	190
19 Unitäre Vektorräume	204
20 Die unitäre Gruppe	209
21 Selbstadjungierte Operatoren	215
22 Quadratische Formen	223

Vorwort

Dies ist das Skriptum der vierstündigen Anfängervorlesung im Wintersemester 2004/05. Im wesentlichen ist es eine korrigierte Neuauflage der Vorlesungsausarbeitung aus dem WS 03/04. Einige Tippfehler und Unklarheiten wurden beseitigt.

Die Vorlesung richtet sich an Studentinnen und Studenten folgender Studiengänge:

- Diplom Mathematik
- Diplom Physik
- Diplom Angewandte Systemwissenschaft
- Bachelor Mathematik/Informatik
- Bachelor Physik mit Informatik
- Bachelor Cognitive Science
- Zweifächerbachelor mit Fach Mathematik
- Lehramt an Gymnasien mit Fach Mathematik
- Lehramt an berufsbildenden Schulen mit Fach Mathematik
- Magister mit Fach Mathematik
- Mathematik als Nebenfach anderer Diplom-, Master- oder Bachelor-Studiengänge

1 Lösen von Gleichungen

Die Algebra beschäftigte sich früher ausschließlich mit dem Lösen algebraischer Gleichungen. Doch im neunzehnten Jahrhundert vollzog sich nach und nach ein tiefgreifender Wandel. Abstrakte algebraische Strukturen waren jetzt der Gegenstand der Untersuchung. Heute durchdringen algebraische Methoden und algebraische Terminologie fast alle Bereiche der Mathematik.

Bevor wir mit dem Studium der linearen Algebra und der grundlegenden algebraischen Strukturen beginnen, wollen wir uns einige einfache Beispiele näher ansehen. Die einfachsten Beispiele algebraischer Gleichungen sind:

$$ax + b = 0, \quad (1)$$

$$ax^2 + bx + c = 0, \quad (2)$$

$$ax^3 + bx^2 + cx + d = 0. \quad (3)$$

Hier sind a, b, c, d fest vorgegebene Zahlen und x ist die *Unbekannte*. Ersetzt man x durch eine Zahl, sagen wir x_0 , so kann man prüfen, ob die vorgelegte Gleichung für diesen Wert erfüllt ist. Ist das der Fall, so nennt man x_0 eine Lösung der Gleichung. Eine wesentliche Aufgabe der Algebra ist es, die Lösbarkeit von Gleichungen zu untersuchen und Lösungsverfahren zu entwickeln.

(1) ist eine lineare Gleichung. Sie besitzt die Lösung

$$x = -\frac{b}{a}.$$

Sind a, b rationale Zahlen, $a \neq 0$, so ist die Lösung x ebenfalls eine rationale Zahl. Gleichung (2) kann man bekanntlich folgendermaßen lösen (quadratische Ergänzung):

$$\begin{aligned} 0 &= x^2 + \frac{b}{a}x + \frac{c}{a} = x^2 + \frac{b}{a}x + \left(\frac{b}{2a}\right)^2 - \left(\frac{b}{2a}\right)^2 + \frac{c}{a} = \left(x + \frac{b}{2a}\right)^2 - \left(\left(\frac{b}{2a}\right)^2 - \frac{c}{a}\right) \\ &= \left(x + \frac{b}{2a}\right)^2 - \frac{1}{(2a)^2}(b^2 - 4ac), \end{aligned}$$

also

$$\left(x + \frac{b}{2a}\right)^2 = \frac{1}{(2a)^2}(b^2 - 4ac)$$

und somit

$$x + \frac{b}{2a} = \pm \frac{1}{2a}\sqrt{b^2 - 4ac}.$$

Die quadratische Gleichung (2) hat also die Lösungen

$$x_1 = \frac{-b + \sqrt{b^2 - 4ac}}{2a}, \quad x_2 = \frac{-b - \sqrt{b^2 - 4ac}}{2a}.$$

Natürlich müssen hier a, b, c so gewählt sein, dass die Terme einen Sinn ergeben.

Sind a, b, c reelle Zahlen und ist $a \neq 0$ und $b^2 - 4ac > 0$, so erhält man zwei verschiedene reelle Lösungen. Ist $b^2 - 4ac = 0$, so fallen beide Lösungen zu einer Lösung zusammen. Ist dagegen $b^2 - 4ac < 0$, so sind die Lösungen als komplexe Zahlen zu interpretieren.

Wie steht es nun mit der Gleichung (3), der sogenannten *kubischen Gleichung*? Der Einfachheit halber sei $a = 1$. Wir untersuchen also die Gleichung

$$x^3 + bx^2 + cx + d = 0. \quad (4)$$

Die beiden ersten Terme kommen in der binomischen Formel

$$\left(x + \frac{b}{3}\right)^3 = x^3 + bx^2 + \frac{b^2}{3}x + \frac{b^3}{27}$$

vor. Man kann also die Gleichung (4) vereinfachen, indem man die Substitution $t = x + \frac{b}{3}$ durchführt. Zunächst erhält man, indem man in (4) den Term $x^3 + bx^2$ durch den Term $t^3 - \frac{b^2}{3}x - \frac{b^3}{27}$ ersetzt, die Gleichung

$$t^3 - \frac{b^2}{3}x - \frac{b^3}{27} + cx + d = 0.$$

Ersetzt man hier noch x durch $t - \frac{b}{3}$ und ordnet nach Potenzen von t , so ergibt sich die einfachere Gleichung

$$t^3 + \left(c - \frac{b^2}{3}\right)t + d - \frac{cb}{3} + \frac{b^3}{9} - \frac{b^3}{27} = 0.$$

Wir haben damit (3) auf eine kubische Gleichung vom Typ

$$x^3 + px - q = 0 \quad (5)$$

reduziert. Der quadratische Term ist verschwunden! Die kubische Gleichung (5) wurde schon in der italienischen Renaissance gelöst (Cardano, Ferro, Tartaglia, 1545).

Schauen wir uns zunächst einmal den Spezialfall $p = 0$, $q = 1$, also die Gleichung

$$x^3 - 1 = 0 \quad (6)$$

an. Welches sind die Lösungen? Die **dritten Einheitswurzeln** $\sqrt[3]{1}$, also (eventuell auch komplexe) Zahlen x mit der Eigenschaft $x^3 = 1$. Auf dem Einheitskreis betrachten wir den Punkt ω zum Winkel $\frac{2\pi}{3} = 120^\circ$.

$$\begin{array}{cccc} & & \omega & \\ & & & \frac{2\pi}{3} \\ -1 & -\frac{1}{2} & 0 & 1 \end{array}$$

$$\omega^2$$

Da $\cos \frac{2\pi}{3} = -\frac{1}{2}$ ist, ist $\sin \frac{2\pi}{3} = \sqrt{1 - \cos^2 \left(\frac{2\pi}{3}\right)} = \sqrt{1 - \frac{1}{4}} = \frac{\sqrt{3}}{2}$, und somit ist ω als komplexe Zahl betrachtet

$$\omega = \frac{-1 + \sqrt{-3}}{2} = \frac{-1 + \sqrt{3} \cdot \sqrt{-1}}{2}.$$

Nach den üblichen Rechenregeln erhält man

$$\begin{aligned} \omega^2 &= \frac{1}{4} (1 - 2\sqrt{-3} + (-3)) = \frac{-1 - \sqrt{-3}}{2}, \\ \omega^3 &= \omega^2 \omega = \frac{1}{4} (1 - \sqrt{-3}^2) = \frac{1}{4} (1 + 3) = 1. \end{aligned}$$

Die Gleichung (6) hat die drei verschiedenen Lösungen

$$x_1 = \omega, \quad x_2 = \omega^2, \quad x_3 = 1.$$

Es gilt also

$$x^3 - 1 = (x - \omega)(x - \omega^2)(x - 1),$$

und daraus folgt durch Ausmultiplizieren und Koeffizientenvergleich

$$\omega + \omega^2 + 1 = 0. \tag{7}$$

Dies findet man auch schön im Bild 1 bestätigt, wenn man ω , ω^2 als Vektoren auffasst. Der Summenvektor $\omega + \omega^2$ ist in der Tat der "Vektor" -1 .

Die Lösung der allgemeinen Gleichung (5) kann man so beschreiben (Eine Herleitung wollen wir hier nicht geben, siehe [3], Seite 621 ff.): Setze $D = -4p^3 - 27q^2$, und wähle dritte Wurzeln

$$u = \sqrt[3]{\frac{27}{2}q + \frac{3}{2}\sqrt{-3D}}, \quad v = \sqrt[3]{\frac{27}{2}q - \frac{3}{2}\sqrt{-3D}}.$$

Dann ist $uv = -3p\sqrt[3]{1}$, also $uv = -3p$, $-3p\omega$ oder $-3p\omega^2$. u und v sollen nun so gewählt sein, dass $uv = -3p$ gilt. Dann sind

$$x_1 = \frac{1}{3}(u + v), \quad x_2 = \frac{1}{3}(\omega^2 u + \omega v), \quad x_3 = \frac{1}{3}(\omega u + \omega^2 v) \tag{8}$$

die Lösungen von (5).

Wer will, kann die Probe machen! Man errechnet:

$$\begin{aligned}x_1 + x_2 + x_3 &= 0, \\x_1x_2 + x_1x_3 + x_2x_3 &= p, \\x_1x_2x_3 &= q.\end{aligned}$$

Diese drei Gleichungen sind äquivalent zu

$$x^3 + px - q = (x - x_1)(x - x_2)(x - x_3).$$

Auch die allgemeine Gleichung vierten Grades lässt sich durch "Radikale" lösen.

Aber es dauerte fast dreihundert Jahre, bis den Versuchen auch die allgemeine Gleichung n -ten Grades

$$x^n + a_1x^{n-1} + \dots + a_{n-1}x + a_n = 0 \quad (9)$$

für $n > 4$ durch Radikale zu lösen ein Ende bereitet wurde. Es gilt nämlich der berühmte

Satz 1.1 (*Abel*, 1827): Die allgemeine Gleichung n -ten Grades ist nicht durch Radikale lösbar, wenn $n \geq 5$ ist.

Die Untersuchung der Gleichung (9) für konkret gegebene rationale Koeffizienten a_1, \dots, a_n ist Gegenstand der so genannten Galois-Theorie und wird in Vorlesungen über Algebra behandelt.

Beispiel 1.2 Die Gleichung

$$x^5 - 1 = 0$$

hat neben $x = 1$ die Lösungen der Gleichung

$$x^4 + x^3 + x^2 + x + 1 = 0.$$

$x^5 - 1 = 0$ ist durch Radikale lösbar. Die vier von 1 verschiedenen Lösungen sind

$$\begin{aligned}\alpha &= -\frac{1}{4} + \frac{\sqrt{5}}{4} + \frac{i}{2}\sqrt{\frac{1}{2}(5 + \sqrt{5})} \\ \beta &= -\frac{1}{4} - \frac{\sqrt{5}}{4} - \frac{i}{2}\sqrt{\frac{1}{2}(5 - \sqrt{5})}\end{aligned}$$

und die dazu konjugiert komplexen Zahlen.

Man sieht an der Gestalt der Lösungen, dass man das regelmäßige Fünfeck mit Zirkel und Lineal konstruieren kann. Wie?

Beispiel 1.3 Dagegen ist die Gleichung

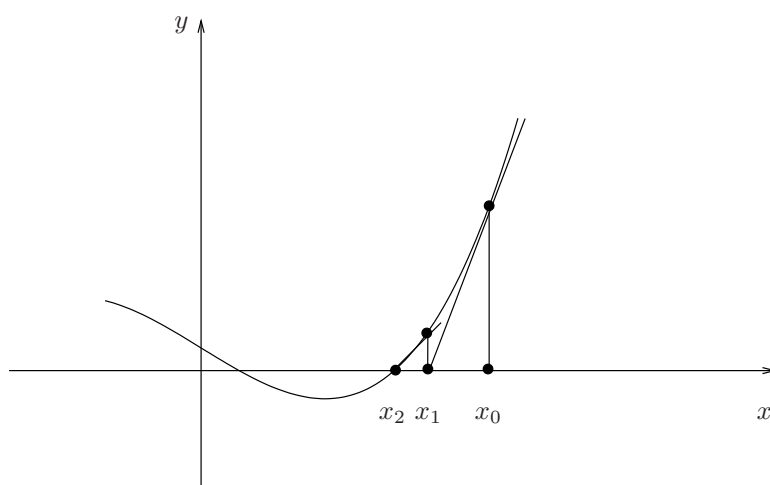
$$x^5 - x - 1 = 0$$

nicht durch Radikale lösbar.

Es sei aber bemerkt, dass man stets Lösungen der Gleichung (9) durch analytische Methoden gewinnen kann. Wir illustrieren dies an der Gleichung

$$x^5 - x - 1 = 0.$$

Offenbar geht es darum, die Nullstellen der Funktion $f(x) = x^5 - x - 1$ zu finden. Da $\lim_{x \rightarrow +\infty} f(x) = +\infty$ und $\lim_{x \rightarrow -\infty} f(x) = -\infty$ gilt, gibt es sicher eine reelle Nullstelle. Der Graph der Funktion in der x, y -Ebene muss die x -Achse schneiden. Einen solchen Schnittpunkt finden wir durch ein Iterationsverfahren (Newton-Verfahren). Wir starten mit x_0 und ersetzen $f(x)$ durch die Tangenten-Funktion im Punkt $(x_0, f(x_0))$, also durch $T_{x_0}(x) = f(x_0) + f'(x_0) \cdot (x - x_0)$.



$T_{x_0}(x) = 0$ ist eine Gleichung vom Grad 1 in x , also leicht lösbar ($f'(x_0) \neq 0$ sollte gelten!).

$$x_1 = x_0 - \frac{f(x_0)}{f'(x_0)}$$

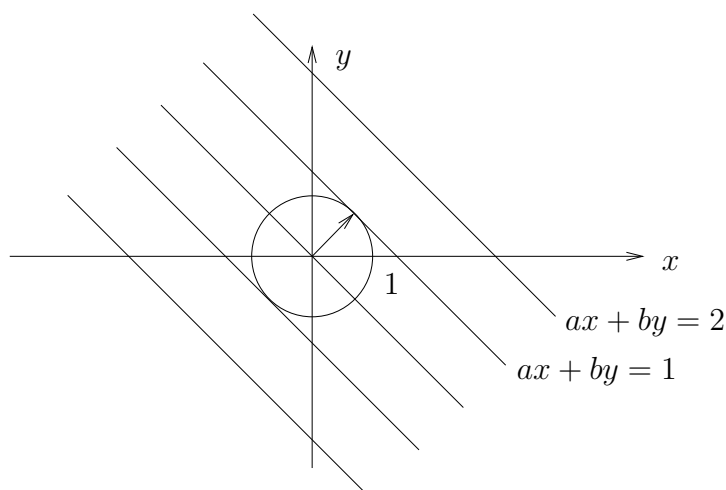
ist die Lösung. Wir wiederholen das Spiel mit x_1 und erhalten

$$x_2 = x_1 - \frac{f(x_1)}{f'(x_1)}.$$

Für $n \geq 0$ ist $x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}$. In unserem Beispiel ist $x_{n+1} = x_n - \frac{x_n^5 - x_n - 1}{5x_n^4 - 1}$. Ist x_0 geeignet gewählt, so konvergiert die Folge x_n gegen eine Nullstelle von f . Startet man etwa mit $x_0 = 1$, so erhält man eine Folge rationaler Zahlen x_n , die gegen eine Nullstelle von $x^5 - x - 1$ konvergiert ($x \approx 1,1673$). Mehr darüber erfährt man in der Analysis.

Bisher haben wir nur Gleichungen mit einer Unbekannten betrachtet. Aber auch das Studium von Gleichungen mit zwei Unbekannten x, y oder, noch allgemeiner, mit n Unbekannten x_1, \dots, x_n ist wichtig, gerade auch im Hinblick auf die Anwendungen innerhalb und außerhalb der Mathematik.

Einige Beispiele algebraischer Gleichungen:



$$ax + by - c = 0. \quad (10)$$

Dies ist die Gleichung einer Geraden in der x, y -Ebene (sofern a und b nicht beide Null sind). Ist $a^2 + b^2 = 1$, so kann man die Zahl $|c|$ deuten als den Abstand der Geraden vom Koordinatenursprung. Das Vorzeichen von c gibt an, ob die Gerade auf der Seite von (a, b) liegt.

Ein Zahlenpaar (x_0, y_0) heißt Lösung von (10), wenn die Gleichung für $x = x_0, y = y_0$ erfüllt ist, d.h. wenn $ax_0 + by_0 - c$ den Wert Null hat. Hat $ax_0 + by_0 - c$ nicht den Wert Null, so ist (x_0, y_0) keine Lösung von (10). Man kann leicht prüfen, ob ein vorgelegtes Paar (x_0, y_0) eine Lösung von (8) ist: Man berechnet einfach den Wert $ax_0 + by_0 - c$.

Die Gleichung (10) hat unendlich viele Lösungen (immer vorausgesetzt, dass a und b nicht beide Null sind). Ist $b \neq 0$, so gibt es zu jedem Wert x_0 genau einen Wert y_0 , so dass (x_0, y_0) eine Lösung von (10) ist:

$$y_0 = -\frac{a}{b}x_0 + \frac{c}{b}.$$

Ist $a \neq 0$, so gibt es zu jedem Wert y_0 genau einen Wert x_0 , so dass (x_0, y_0) Lösung von (10) ist:

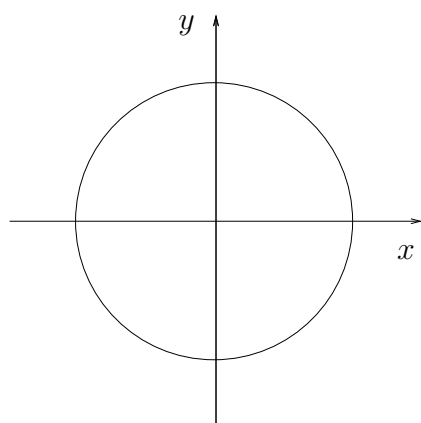
$$x_0 = -\frac{b}{a}y_0 + \frac{c}{a}.$$

Komplizierter ist die Gleichung zweiten Grades

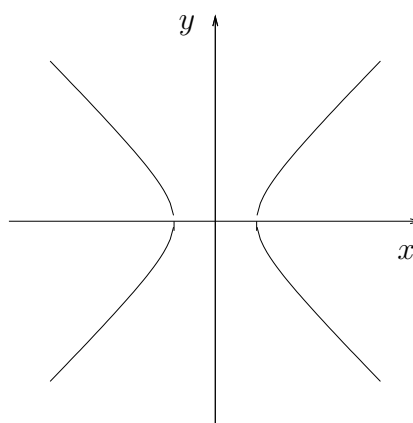
$$ax^2 + bxy + cy^2 + dx + ey + f = 0. \quad (11)$$

Wieder ist die Lösungsmenge eine geometrische Figur in der x, y -Ebene, diesmal nicht linear, sondern ein "Kegelschnitt" (falls a, b, c nicht alle gleich Null sind). Ein-

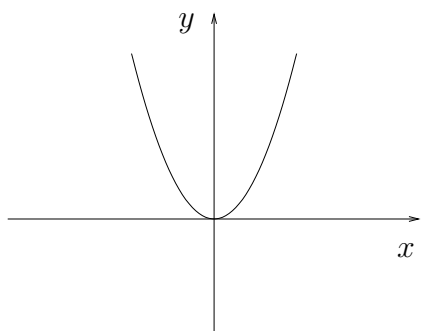
fache Beispiele sind:



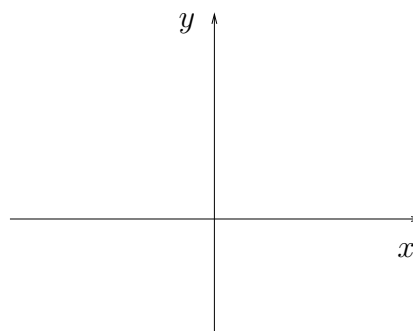
Kreis $x^2 + y^2 = 1$



Hyperbel $x^2 - y^2 = 1$



Parabel $x^2 - y = 0$



Achsenkreuz $xy = 0$

Man sieht: Durch die Einführung von Koordinaten zur Beschreibung von Punkten, hier (x, y) , für Punkte in der euklidischen Ebene, sind die Gebilde der Geometrie einer algebraischen Behandlung zugänglich geworden. Sie werden durch algebraische Gleichungen beschrieben.

Dieses Gebiet der Mathematik nennt man in klassischer Terminologie *analytische Geometrie* und in moderner Sprache *algebraische Geometrie*.

Ein weiterer wesentlicher Schritt der Verallgemeinerung, den wir durchführen müssen, besteht darin, an Stelle einzelner Gleichungen Gleichungssysteme in mehreren Unbekannten zu behandeln. Dabei ist es auch wichtig, als Unbekannte nicht nur Zahlen, sondern auch andere Typen von Variablen, wie Funktionen, zuzulassen. In der modernen Mathematik sieht man daher ganz von der spezifischen Eigenart der Elemente, die in den Gleichungen vorkommen, ab, schafft dann, ausgehend von dem konkreten Problem eine passende abstrakte mathematische Theorie, die dann nicht nur auf das Ausgangsproblem, sondern auf viele verschiedene konkrete Modelle dieser mathematischen Struktur anwendbar ist. So ist zum Beispiel die Theorie der Vektorräume der Rahmen für die Behandlung *linearer Probleme*. Entstanden aus den linearen Gleichungssystemen für reelle Variable x_1, \dots, x_n , kann man die Theorie auch auf lineare Differenzialgleichungssysteme und Integralgleichungen der mathematischen Systemtheorie anwenden.

Es ist ein Ziel der Vorlesung, diesen modernen Aspekt der Mathematik, den man

auch mit Schlagwörtern *axiomatische Methode* oder *mathematische Strukturen* belegt, am Beispiel der Algebra klar herauszuarbeiten.

Bevor wir uns der abstrakten Theorie zuwenden, wollen wir einen ersten Versuch zu reellen linearen Gleichungssystemen unternehmen.

2 Lineare Gleichungssysteme, erste Schritte

Das einfachste Beispiel

$$ax = b \quad (12)$$

haben wir schon untersucht. Ist $a \neq 0$, so hat (12) genau eine Lösung, nämlich $x = \frac{b}{a}$. Ist $a = 0$ und $b \neq 0$, so ist (12) nicht lösbar, d.h. es gibt keine Zahl x , so dass (12) erfüllt ist, denn für alle x gilt

$$ax = 0x = 0 \neq b.$$

Ist $a = 0$ und $b = 0$, so ist jeder Wert x eine Lösung, denn es gilt ja $ax = 0x = 0 = b$ für alle x .

Kommen wir zum Fall eines linearen Gleichungssystems mit zwei Gleichungen und zwei Unbekannten x, y . Es sind Konstanten a, b, c, d, e, f gegeben. Das Gleichungssystem hat die Form

$$\begin{aligned} ax + by &= e \\ cx + dy &= f \end{aligned} \quad (13)$$

Um aus dem Gleichungssystem y zu eliminieren, multiplizieren wir die erste Gleichung von (13) mit d und die zweite Gleichung mit b . Wir erhalten:

$$\begin{aligned} adx + bdy &= de \\ bcx + bdy &= bf \end{aligned} \quad (14)$$

Subtrahieren wir nun von der ersten Gleichung in (14) die zweite Gleichung, so ergibt sich

$$adx - bcx = de - bf$$

und somit

$$(ad - bc)x = de - bf. \quad (15)$$

Analog verfahren wir, um aus (13) x zu eliminieren. Dieses Mal erhalten wir aus (13)

$$\begin{aligned} acx + bcy &= ce \\ acx + ady &= af \end{aligned} \quad (16)$$

und durch Subtraktion der ersten von der zweiten Gleichung ergibt sich

$$(ad - bc)y = af - ce. \quad (17)$$

Nehmen wir nun an, dass die Konstanten a, b, c, d die Ungleichung $ad - bc \neq 0$ erfüllen. Dann erhalten wir aus (15) und (17) explizit:

$$\begin{aligned} x &= \frac{de - bf}{ad - bc} \\ y &= \frac{af - ce}{ad - bc} \end{aligned} \quad (18)$$

Durch Einsetzen von (18) in (13) sieht man, dass diese Werte x, y das Gleichungssystem (13) erfüllen:

$$\begin{aligned} ax + by &\stackrel{(18)}{=} \frac{a(de - bf) + b(af - ce)}{ad - bc} \\ &= \frac{ade - abf + abf - bce}{ad - bc} = e. \end{aligned}$$

und genauso $cx + dy = f$. Damit haben wir gezeigt:

Lemma 2.1 (Cramersche Regel): Ist $ad - bc \neq 0$, so ist

$$(x, y) = \left(\frac{de - bf}{ad - bc}, \frac{af - ce}{ad - bc} \right)$$

die einzige Lösung von (13).

Aus diesem Grund heißt die Zahl $ad - bc$ auch die **Determinante** des linearen Gleichungssystems (13). Man schreibt dafür

$$\begin{vmatrix} a & b \\ c & d \end{vmatrix} = \det \begin{pmatrix} a & b \\ c & d \end{pmatrix} = ad - bc.$$

Ist $ad - bc = 0$, so kann es unendlich viele oder gar keine Lösungen geben.

Beispiel 2.2

a)

$$\begin{cases} x + y &= 1 \\ 2x + 2y &= 3 \end{cases}$$

hat die Determinante $\begin{vmatrix} 1 & 1 \\ 2 & 2 \end{vmatrix} = 0$ und keine Lösungen. Wäre (x, y) eine Lösung, so wäre $2x + 2y = 2$ und $2x + 2y = 3$, also $2 = 3$, was Unsinn ist.

b)

$$\begin{cases} x + 2y &= 3 \\ 2x + 4y &= 6 \end{cases}$$

hat die Determinante $\begin{vmatrix} 1 & 2 \\ 2 & 4 \end{vmatrix} = 0$ und unendlich viele Lösungen, nämlich $(x, -\frac{1}{2}x + \frac{3}{2})$, x beliebig.

Die beiden Gleichungen des Gleichungssystems sind äquivalent.

Man kann also eine der beiden weglassen, ohne die Lösungsmenge zu ändern. Damit wird klar, dass die Lösungsmenge eine Gerade in der (x, y) -Ebene ist. Sie hat die "Gleichung" $y = -\frac{1}{2}x + \frac{3}{2}$.

Der allgemeinste Typ eines linearen Gleichungssystems hat die folgende Form:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n = b_2 \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n = b_m \end{cases} \quad (19)$$

Hier sind m, n beliebige natürliche Zahlen:

$$\begin{aligned} m &= \text{Anzahl der Gleichungen,} \\ n &= \text{Anzahl der Unbekannten.} \end{aligned}$$

x_1, \dots, x_n sind die Unbekannten des Gleichungssystems.

Die $m \cdot n$ reellen Zahlen a_{ij} heißen die **Koeffizienten** und die m reellen Zahlen b_i heißen die **konstanten Terme** des Gleichungssystems.

Das Gleichungssystem (19) heißt **homogen**, wenn $b_i = 0$ für $i = 1, \dots, m$ gilt. (19) heißt **inhomogen**, wenn für mindestens ein i der konstante Term b_i ungleich Null ist.

Das lineare Gleichungssystem

$$\begin{cases} a_{11}x_1 + \dots + a_{1n}x_n = 0 \\ \vdots \\ a_{m1}x_1 + \dots + a_{mn}x_n = 0 \end{cases} \quad (20)$$

heißt das zu (19) gehörige **homogene lineare Gleichungssystem**.

Die Koeffizienten des Gleichungssystems werden folgendermaßen in einer Tabelle angeordnet

$$\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix} \quad (21)$$

Solch eine Tabelle heißt $m \times n$ -**Matrix**. (21) ist die Matrix des linearen Gleichungssystems (19). Abkürzend schreibt man für die Matrix (21) auch

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

oder noch kürzer einfach (a_{ij}) , wenn klar ist, in welchem Bereich die Indizes i, j laufen. Oft ist es zweckmäßig, der Matrix einen Namen zu geben, etwa $A = (a_{ij})$.

Für $i = 1, \dots, m$ heißt

$$(a_{i1}, \dots, a_{in})$$

die i -te **Zeile** von A , und für $j = 1, \dots, n$ heißt

$$\begin{pmatrix} a_{1j} \\ \vdots \\ a_{mj} \end{pmatrix}$$

die j -te **Spalte** von A . Um Platz zu sparen, aber andererseits deutlich zwischen Spalten und Zeilen zu unterscheiden, schreiben wir auch

$$[a_{1j}, \dots, a_{mj}] \quad \text{für die Spalte} \quad \begin{pmatrix} a_{1j} \\ \vdots \\ a_{mj} \end{pmatrix}.$$

Die eckigen Klammern sollen andeuten, dass wir eine Spalte meinen.

Beispiel 2.3 $[a, b, c] = \begin{pmatrix} a \\ b \\ c \end{pmatrix} \neq (a, b, c)$.

Fügt man die Spalte

$$b = [b_1, \dots, b_m]$$

der konstanten Terme des Gleichungssystems (19) zur Matrix A hinzu, so erhält man eine $m \times (n + 1)$ -Matrix

$$(A, b) = (a_{ij}, b_i) = \left(\begin{array}{ccc|c} a_{11} & \cdots & a_{1n} & b_1 \\ \vdots & & \vdots & \vdots \\ a_{m1} & \cdots & a_{mn} & b_m \end{array} \right)$$

Diese Matrix heißt die **erweiterte Matrix** des linearen Gleichungssystems (19).

Definition 2.4 Ein n -Tupel $x_0 = [x_1^0, \dots, x_n^0]$ reeller Zahlen heißt **Lösung** des linearen Gleichungssystems (19), wenn

$$\begin{aligned} a_{11}x_1^0 + \dots + a_{1n}x_n^0 &= b_1 \\ \vdots & \\ a_{m1}x_1^0 + \dots + a_{mn}x_n^0 &= b_m \end{aligned}$$

gilt. x_i^0 heißt **i -te Komponente** der Lösung x_0 .

Das Gleichungssystem (19) heißt

- a) **lösbar**, wenn es mindestens eine Lösung besitzt,
- b) **nicht lösbar**, wenn es keine Lösung besitzt und
- c) **eindeutig lösbar**, wenn es genau eine Lösung besitzt.

Wie bei allen Gleichungen stellen sich nun zwei Probleme:

- I. Wann ist das Gleichungssystem lösbar?
Gibt es einfache notwendige und hinreichende Bedingungen dafür?
- II. Wie sieht die Menge der Lösungen eines lösbaren linearen Gleichungssystems aus?

Es sei ein weiteres lineares Gleichungssystem

$$\begin{aligned} a'_{11}x_1 + \dots + a'_{1n}x_n &= b'_1 \\ \vdots & \\ a'_{k1}x_1 + \dots + a'_{kn}x_n &= b'_k \end{aligned} \tag{22}$$

mit k Gleichungen und n Unbekannten gegeben. Mit (A', b') bezeichnen wir die erweiterte Matrix von (22).

Definition 2.5 Die Gleichungssysteme (19) und (22) heißen **äquivalent** (wir schreiben dann $(19) \sim (22)$ oder auch $(A, b) \sim (A', b')$), wenn sie dieselbe Lösungsmenge besitzen, d.h. wenn gilt:

Ist x_0 eine Lösung von (19), so ist x_0 auch eine Lösung von (22) und ist x_0 eine Lösung von (22), so ist x_0 auch eine Lösung von (19).

Wir betrachten nun **elementare Umformungen** des Gleichungssystems (19).

Umformungen vom Typ I:

T_{ij} : Vertausche die i -te und j -te Gleichung.

Umformungen vom Typ II:

$T_{ij}(c)$: Ersetze die i -te Gleichung durch die Gleichung

$$(a_{i1} + ca_{j1})x_1 + \dots + (a_{in} + ca_{jn})x_n = b_i + cb_j.$$

Hierbei sind i, j Indizes mit $1 \leq i \leq m; 1 \leq j \leq m; i \neq j$ und c ist eine reelle Zahl.

Umformungen vom Typ III:

$T_i(c)$: Ersetze die i -te Gleichung durch die Gleichung

$$ca_{i1}x_1 + \dots + ca_{in}x_n = cb_i.$$

Hierbei ist c eine von Null verschiedene reelle Zahl.

Umformung vom Typ IV:

T_i : Weglassen der i -ten Gleichung, falls diese Null ist, d. h. falls alle Koeffizienten $a_{i,j}, j = 1, \dots, n$ und b_i Null sind.

Beispiel 2.6

$$\begin{array}{l} 2x_1 + x_2 + x_3 = 1 \\ x_1 + 2x_2 + 2x_3 = 2 \\ x_1 + x_2 + x_3 = 0 \end{array} \xrightarrow{T_{12}} \begin{array}{l} x_1 + 2x_2 + 2x_3 = 2 \\ 2x_1 + x_2 + x_3 = 1 \\ x_1 + x_2 + x_3 = 0 \end{array} \xrightarrow{T_{21}(-2)} \begin{array}{l} x_1 + 2x_2 + 2x_3 = 2 \\ x_1 + 2x_2 + 2x_3 = 2 \\ x_1 + x_2 + x_3 = 0 \end{array} \xrightarrow{T_{31}(-1)} \begin{array}{l} x_1 + 2x_2 + 2x_3 = 2 \\ -3x_2 - 3x_3 = -3 \\ x_1 + x_2 + x_3 = 0 \end{array} \xrightarrow{T_{32}(-\frac{1}{3})} \begin{array}{l} x_1 + 2x_2 + 2x_3 = 2 \\ -3x_2 - 3x_3 = -3 \\ -x_2 - x_3 = -2 \end{array}$$

$$\begin{array}{l} x_1 + 2x_2 + 2x_3 = 2 \\ -3x_2 - 3x_3 = -3 \\ 0x_3 = -1 \end{array}$$

Das letzte Gleichungssystem ist nicht lösbar, denn die Gleichung $0x_3 = -1$ ist nicht lösbar. Also ist auch das Ausgangssystem nicht lösbar! Es gilt nämlich

Satz 2.7 Zwei lineare Gleichungssysteme sind äquivalent, wenn das eine aus dem anderen durch eine endliche Folge elementarer Umformungen hervorgegangen ist.

Beweis:

- a) Offensichtlich ändert eine elementare Umformung vom Typ I,III und IV die Lösungsmenge nicht.
- b) Sei (22) aus (19) durch die elementare Umformung $T_{ij}(c)$ entstanden. Ist nun x_0 eine Lösung von (19), so gilt insbesondere

$$\begin{aligned} a_{i1}x_1^0 + \dots + a_{in}x_n^0 &= b_i & \text{und} \\ a_{j1}x_1^0 + \dots + a_{jn}x_n^0 &= b_j, \end{aligned}$$

und somit gilt $(a_{i1} + ca_{j1})x_1^0 + \dots + (a_{in} + ca_{jn})x_n^0 = a_{i1}x_1^0 + \dots + a_{in}x_n^0 + c(a_{j1}x_1^0 + \dots + a_{jn}x_n^0) = b_i + cb_j$,

d.h. x_0 ist auch eine Lösung von (22).

Nun ist aber umgekehrt (19) aus (22) durch die elementare Umformung $T_{ij}(-c)$ entstanden, und deshalb gilt auch: Ist x_0 Lösung von (22), so ist x_0 auch eine Lösung von (19).

- c) Aus a) und b) folgt jetzt der Satz. □

Wir wollen erklären, was ein lineares Gleichungssystem in "Stufenform" ist. Zum Beispiel sind die folgenden Gleichungssysteme in Stufenform

a)

$$\begin{array}{rccccrcr} x_1 & + & 3x_2 & + & 4x_3 & - & x_4 & = & 1 \\ & & x_2 & & & - & 2x_4 & = & 0 \\ & & & & 3x_3 & & & = & 1 \end{array}$$

b)

$$\begin{array}{rccccrcr} x_1 & + & x_2 & + & x_3 & + & x_4 & = & 1 \\ & & & & x_3 & - & x_4 & = & 1 \\ & & & & & & x_4 & = & 1 \end{array}$$

Die erweiterte Matrix ist hier

$$\begin{pmatrix} 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & -1 & 1 \\ 0 & 0 & 0 & 1 & 1 \end{pmatrix}$$

c)

$$\begin{array}{rccccrcr} x_1 & + & & x_4 & = & 2 \\ x_2 & + & & x_3 & = & 1 \\ & & & x_3 & = & 2 \\ & & & 0 \cdot x_4 & = & 1 \end{array}$$

Die erweiterte Matrix hat hier die Form

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

Die vierte Zeile von A ist Null, der vierte Koeffizient von b aber ungleich Null. Dieses Gleichungssystem ist nicht lösbar.

Die allgemeine Definition der Stufenform ist vielleicht zunächst etwas schwierig wegen der auftretenden Indizes. Daran muss man sich gewöhnen!

Definition 2.8 Das lineare Gleichungssystem (22) hat **Stufenform**, wenn es eine Zahl r mit $0 \leq r \leq k$ und Indizes $1 \leq j_1 < j_2 < \dots < j_r \leq n$ gibt, so dass gilt:

$$a'_{ij} = 0 \text{ für } i = 1, \dots, r, j < j_i \text{ und für } i > r, j = 1, \dots, n$$

und

$$a'_{1j_1}, \dots, a'_{rj_r} \neq 0.$$

(22) hat also die Gestalt

$$\begin{array}{rcl} a'_{1j_1}x_{j_1} + \dots & + a'_{1n}x_n & = b'_1 \\ & + a'_{2n}x_n & = b'_2 \\ & \dots & \\ & + a'_{rj_r}x_{j_r} + \dots & + a'_{rn}x_n = b'_r \\ & & 0x_n = b'_{r+1} \\ & & \vdots \\ & & 0x_n = b'_k. \end{array} \quad (23)$$

Für $i = 1, \dots, r$ hängt also die i -te Gleichung nur von Unbekannten x_j mit einer Nummer $j \geq j_i$ ab.

Die erweiterte Matrix von (23) hat die Gestalt

$$\left(\begin{array}{cccc|c} 0 & \dots & 0 & a'_{1j_1} & \dots & a'_{1n} & b'_1 \\ 0 & \dots & 0 & 0 & \dots & a'_{2n} & b'_2 \\ \vdots & & & & & \vdots & \vdots \\ 0 & \dots & & & 0 & a'_{rj_r} & \dots & a'_{rn} & b'_r \\ 0 & \dots & & & & & & 0 & b'_{r+1} \\ \vdots & & & & & & & \vdots & \vdots \\ 0 & \dots & & & & & & 0 & b'_k \end{array} \right)$$

Man sagt dann auch: Die Matrix A' von (23) hat **Zeilenstufenform**.

Der Schlüssel für das Lösen von Gleichungssystemen ist der folgende sehr wichtige Satz.

Satz 2.9 Jedes lineare Gleichungssystem ist zu einem linearen Gleichungssystem in Stufenform äquivalent.

Beweis: Es sei ein lineares Gleichungssystem (19) gegeben. Durch sukzessive Anwendung elementarer Umformungen vom Typ I und II wollen wir dies auf Stufenform bringen. Der Beweis ergibt einen Algorithmus zur Bestimmung einer Stufenform und lässt sich leicht in ein Computerprogramm übersetzen (Übung für Computerfreaks!).

1. Schritt: Sind alle Koeffizienten des Gleichungssystems gleich Null, so liegt schon eine Stufenform vor. Sei nun $a_{ij} \neq 0$ für wenigstens ein Indexpaar. Dann sei j_1 der kleinste Index j ($1 \leq j \leq n$), so dass die j -te Spalte der Matrix A des Gleichungssystems nicht nur aus Nullen besteht (Die Unbekannten x_1, \dots, x_{j_1-1} kommen dann im Gleichungssystem nicht vor, falls $j_1 > 1$ ist.). Durch eine elementare Umformung vom Typ I kann man erreichen, dass $a_{1j_1} \neq 0$ gilt.

Dann wende man für $i = 2, \dots, m$ die elementare Umformung

$$T_{i,1} \left(-\frac{a_{ij_1}}{a_{1j_1}} \right)$$

an. Man erhält ein Gleichungssystem, dessen j_1 -te Spalte die Form $[a_{1j_1}, 0, \dots, 0]$ hat, in dem x_{j_1} also nur in der ersten Gleichung vorkommt.

2. Schritt: Man behandle nun das Gleichungssystem, welches aus der zweiten bis m -ten Gleichung besteht, wie in Schritt 1.

Nach spätestens m Schritten hat man die gewünschte Stufenform erhalten. \square

Dieser Algorithmus heißt **Gaußsches Eliminationsverfahren**. Es ist oft übersichtlicher, das Gaußsche Eliminationsverfahren auf die erweiterte Matrix (A, b) des Gleichungssystems anzuwenden anstatt auf die Gleichungen selbst.

Beispiel 2.10

$$\begin{aligned} x_1 + x_2 + x_3 + x_4 &= 1 \\ 2x_1 + x_2 - 2x_3 + 2x_4 &= 2 \\ 4x_1 - 3x_2 + 2x_3 &= 30 \end{aligned}$$

Die erweiterte Matrix ist

$$\left(\begin{array}{cccc|c} 1 & 1 & 1 & 1 & 1 \\ 2 & 1 & -2 & 2 & 2 \\ 4 & -3 & 2 & 0 & 30 \end{array} \right).$$

Das im obigen Beweis beschriebene Verfahren ergibt:

$$\xrightarrow{T_{21}(-2)} \left(\begin{array}{cccc|c} 1 & 1 & 1 & 1 & 1 \\ 0 & -1 & -4 & 0 & 0 \\ 4 & -3 & 2 & 0 & 30 \end{array} \right) \xrightarrow{T_{31}(-4)} \left(\begin{array}{cccc|c} 1 & 1 & 1 & 1 & 1 \\ 0 & -1 & -4 & 0 & 0 \\ 0 & -7 & -2 & -4 & 26 \end{array} \right).$$

Der erste Schritt ist abgeschlossen. Jetzt werden die zweite und die dritte Zeile behandelt:

$$\xrightarrow{T_{32}(-7)} \left(\begin{array}{cccc|c} 1 & 1 & 1 & 1 & 1 \\ 0 & -1 & -4 & 0 & 0 \\ 0 & 0 & 26 & -4 & 26 \end{array} \right).$$

Wir haben die Stufenform

$$\begin{array}{rccccrcr} x_1 & + & x_2 & + & x_3 & + & x_4 & = & 1 \\ & & - & x_2 & - & 4x_3 & & = & 0 \\ & & & & 26x_3 & - & 4x_4 & = & 26 \end{array}$$

erhalten.

Definition 2.11 Ist ein Gleichungssystem aus m Gleichungen und n Unbekannten x_1, \dots, x_n in Stufenform (23) gegeben, so heißen die Unbekannten x_{j_1}, \dots, x_{j_r} **Hauptvariable** und die Unbekannten x_j , $j \neq j_1, \dots, j_r$ heißen **freie Variable** von (23).

Ist ein Gleichungssystem in Stufenform, so kann man die Lösungen leicht ablesen. Es sei die Stufenform (23) gegeben. Dann ist dieses Gleichungssystem höchstens dann lösbar, wenn $b'_{r+1} = \dots = b'_k = 0$ gilt. Ist dies der Fall, so kann man für die freien Variablen x_j ($j \neq j_1, \dots, j_r$) irgendwelche Werte x_j^0 vorgeben. Durch die r -te Gleichung liegt dann die Hauptvariable x_{j_r} fest:

$$x_{j_r} = x_{j_r}^0 := \frac{1}{a'_{rj_r}} (b'_r - a'_{rj_{r+1}} x_{j_{r+1}}^0 - \dots - a'_{rn} x_n^0).$$

Aus der $(r-1)$ -ten Gleichung ergibt sich die Hauptvariable $x_{j_{r-1}}$ usw.. Ist $b'_{r+1} = \dots = b'_k = 0$, so hat das System also $n-r$ **unabhängige** Lösungen, das soll heißen, dass man für die $n-r$ freien Variablen beliebige Werte einsetzen darf. Die Hauptvariablen sind dann festgelegt. Wir haben also:

Satz 2.12 Ein lineares Gleichungssystem (19) hat dieselben Lösungen wie eine mittels des Gaußschen Eliminationsverfahrens gewonnene Stufenform. Hat diese Stufenform die Gestalt (23), so ist das Gleichungssystem (23) genau dann lösbar, wenn $b'_{r+1} = \dots = b'_k = 0$ gilt. In diesem Fall gilt: Zu jeder beliebigen Wahl α_j , $j \neq j_1, \dots, j_r$ von Werten für die freien Variablen der Stufenform (23) gibt es genau eine Lösung $x_0 = (x_1^0, \dots, x_n^0)$ des Gleichungssystems (19) mit $x_j^0 = \alpha_j$ für $j \neq j_1, \dots, j_r$. Das Gleichungssystem (19) ist genau dann eindeutig lösbar, wenn $r = n$ und $b'_{r+1} = \dots = b'_k = 0$ in der zugehörigen Stufenform (23) gilt (Insbesondere muss dann $m \geq n$ gelten.). \square

Korollar 2.13 Ein lineares Gleichungssystem (19) mit $m = n$ ist genau dann eindeutig lösbar, wenn man es durch elementare Umformungen auf die Gestalt

$$\begin{array}{rccccrcr} a'_{11}x_1 + & \dots & a'_{1n}x_n & = & b'_1 \\ & a'_{22}x_2 + \dots & a'_{2n}x_n & = & b'_2 \\ & & \ddots & & \\ & & & & a'_{nn}x_n & = & b'_n \end{array}$$

bringen kann mit $a'_{ii} \neq 0$ für $i = 1, \dots, n$. \square

Man kann eine Zeilenstufenform (23) noch weiter vereinfachen, indem man die Umformungen $T_i(a_{ij_i}^{-1})$ für $i = 1, \dots, r$ anwendet. Dann hat die i -te Gleichung ($i \leq r$) die Form

$$x_{j_i} + a'_{ij_i+1} x_{j_i+1} + \dots + a'_{in} x_n = b'_i$$

Jetzt kann man weiter aus der ersten Gleichung die Unbekannte x_{j_2} eliminieren, indem man das a_{1j_2} -fache der zweiten Gleichung von der ersten Gleichung abzieht. Schließlich kann man die Unbekannte x_{j_r} aus den ersten $r - 1$ Gleichungen eliminieren. Weiter kann man die Gleichungen $0 = 0$ weglassen. Es bleiben genau r Gleichungen übrig, wenn das Gleichungssystem lösbar ist. Wir erhalten so eine sogenannte *reduzierte Zeilenstufenform* des Gleichungssystems. In der ersten Gleichung kommen x_{j_1} und höchstens noch die $n - r$ freien Variablen $x_i, i \neq j_\nu$ vor.

$$x_{j_1} + \dots + a_{1j_2-1}x_{j_2-1} + a_{1j_2+1}x_{j_2+1} + \dots = b_1.$$

Ist das Gleichungssystem nicht lösbar, so hat die reduzierte Zeilenstufenform eine widersprüchliche $(r + 1)$ -te Gleichung $0 = 1$.

Ein Beispiel der erweiterten Matrix einer reduzierten Zeilenstufenform:

$$\begin{pmatrix} 0 & 1 & 7 & 0 & 0 & 23 & 0 & 87 & 0 & 11 \\ 0 & 0 & 0 & 1 & 0 & 6 & 0 & 4 & 0 & 3 \\ 0 & 0 & 0 & 0 & 1 & -5 & 0 & 91 & 0 & 8 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 5 & 0 & 6 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 4 \end{pmatrix}$$

Bei einem lösbaeren linearen Gleichungssystem mit n Gleichungen in n Unbekannten kann man an der reduzierten Zeilenstufenform direkt die Lösung ablesen.

Übungen

1. Es seien a, b reelle Zahlen. Für welche a, b ist das lineare Gleichungssystem

$$\begin{aligned} ax + by &= b \\ x + y &= a \end{aligned}$$

nicht lösbar, lösbar, eindeutig lösbar?

2. Beweisen Sie: Ein homogenes lineares Gleichungssystem mit m Gleichungen und n Unbekannten besitzt unendlich viele Lösungen, falls $m < n$ gilt.
3. Lösen Sie die folgenden linearen Gleichungssysteme:

$$\begin{aligned} & 2x_1 + x_2 - 3x_3 = 1 \\ \text{a) } & x_1 - x_2 - x_3 = 0 \\ & 2x_1 + x_2 + x_3 = 2 \\ & x_1 + 2x_2 + x_3 = 1 \\ \text{b) } & 5x_1 + 4x_2 + 3x_3 = 6 \\ & 2x_1 + x_2 + x_3 = 0 \\ & 2x_1 + 4x_2 + 2x_3 = 2 \\ \text{c) } & x_1 + 2x_2 + 2x_3 + 2x_4 + x_5 = 3 \\ & 2x_1 + 4x_2 + 4x_3 + 4x_4 + 5x_5 = 5 \\ & x_1 + 4x_2 + 3x_3 + 2x_4 + x_5 = 4 \end{aligned}$$

$$\begin{array}{r} x_1 + x_2 + x_3 + x_4 = 0 \\ \text{d) } 5x_1 + 7x_2 + x_3 - x_4 = 2 \\ x_1 + 2x_2 - x_3 - 2x_4 = 1 \end{array}$$

4. Für welche a, b, c ist das lineare Gleichungssystem

$$\begin{array}{r} 3x_1 + 3x_2 + 3x_3 = a \\ 4x_1 + 6x_2 + 2x_3 = b \\ x_1 + 2x_2 = c \end{array}$$

lösbar?

5. Wie findet man eine lineare Gleichung $ax + by + c = 0$, die zwei gegebene Punkte $(x_0, y_0), (x_1, y_1)$ als Lösungen besitzt? Wie viele Möglichkeiten gibt es für a, b, c ?
6. Welcher Bedingung müssen die Zahlen a, b genügen, damit das folgende Gleichungssystem lösbar wird?

$$\begin{array}{r} bx_1 + x_2 + (ab + a + b)x_3 + (ab^2 + b^2)x_4 = a^2 \\ -ax_1 + (a + 1)x_2 + (ab + a + b)x_3 + (ab^2 - a^2b + b^2)x_4 = a^2 + ab^2 \\ (b + 1)x_2 + (b^2 + ab + a + b)x_3 + (b^3 + b^2)x_4 = b^3 + a^2 \end{array}$$

7. Lösen Sie das (nicht lineare) Gleichungssystem

$$\begin{array}{r} x^2 + y^2 = 1 \\ ax + by = 1, \end{array}$$

wobei a, b gegebene reelle Zahlen sind, die nicht beide Null sind.

8. Geben Sie die reelle Lösung der kubischen Gleichung $x^3 + 2x - 1 = 0$ an.

3 Mengen und Abbildungen

Zur präzisen Beschreibung mathematischer Sachverhalte bedient man sich der Mengen-Schreibweise oder der *Mengensprache*. Wir müssen den *Grundwortschatz* dieser Sprache lernen. Wichtige Vokabeln sind:

Menge, Element, Teilmenge, Produkt, Vereinigung, Durchschnitt, Abbildung, injektiv, surjektiv, bijektiv, Bild, Urbild, Umkehrabbildung. Es ist also einiges zu tun.

Definition 3.1 (heuristische Definition, Cantor, 1845 - 1918): Unter einer **Menge** M verstehen wir “eine Zusammenfassung von bestimmten wohl unterschiedenen Objekten unserer Anschauung oder unseres Denkens (welche die **Elemente** der Menge genannt werden) zu einem Ganzen”.

$x \in M$ bedeutet: x ist Element von M .

$x \notin M$ bedeutet: x ist nicht Element von M .

(siehe [25])

Beispiel 3.2

a) $M = \{1, 37\}$ ist die Menge mit den Elementen 1 und 37. Für jedes ‘Element’ x mit $x \neq 1$ und $x \neq 37$ gilt also $x \notin M$.

b) $M = \{1, \{1\}\}$ enthält die verschiedenen Elemente 1 und $\{1\}$.

c) $M = \{1, 2, 4, 8\}$ ist die Menge der Zweierpotenzen zwischen 1 und 10. Statt die Elemente aufzuzählen, kann man auch schreiben $M = \{2^n | n = 0, 1, 2, 3\}$ oder $M = \{x | \exists n \in \{0, 1, 2, 3\} : x = 2^n\}$.

Das Zeichen “ \exists ” bedeutet “*Es existiert*”.

Der Doppelpunkt “:” ist hier eine Abkürzung für “so dass”.

d) Wir benutzen die folgenden Standardbezeichnungen:

$\mathbb{N} = \{0, 1, 2, 3, \dots\}$ ist die Menge der **natürlichen Zahlen** (einschließlich der Null!).

$\mathbb{Z} = \{\dots - 2, -1, 0, 1, 2, \dots\}$ ist die Menge der **ganzen Zahlen**.

$\mathbb{Q} = \{\frac{a}{b} | a \in \mathbb{Z}, b \in \mathbb{Z}, b \neq 0\}$ ist die Menge der **rationalen Zahlen**.

\mathbb{R} ist die Menge aller **reellen Zahlen**.

(Die keinesfalls triviale Konstruktion der reellen Zahlen soll hier nicht gegeben werden, siehe [8].)

Definition 3.3 Für zwei Mengen M, N gilt: N ist **Teilmenge** von M (in Zeichen: $N \subset M$) genau dann, wenn jedes Element x von N auch ein Element von M ist. In Kurzform lautet die Definition so:

$$N \subset M :\Leftrightarrow (\forall x : x \in N \Rightarrow x \in M).$$

Das Zeichen “ $:\Leftrightarrow$ ” bedeutet hier “*definitionsgemäß genau dann, wenn gilt*”.

Das Symbol “ \forall ” bedeutet “*für alle*”. Die Aussage in Klammern heißt also:

Für alle x gilt: Ist $x \in N$, so ist $x \in M$.

Nachdem wir die Enthaltensein-Relation für Mengen eingeführt haben, können wir ein kleines Lemma beweisen.

Lemma 3.4 Es seien A, B, C Mengen. Dann gilt:

- a) $A \subset A$.
- b) Ist $A \subset B$ und $B \subset C$, so ist $A \subset C$.
- c) Ist $A \subset B$ und $B \subset A$, so ist $A = B$.

Beweis: zu a): Für alle x ist die Implikation $(x \in A \Rightarrow x \in A)$ richtig, d.h. es gilt $A \subset A$.

zu b): Wenn für alle x die Implikationen $(x \in A \Rightarrow x \in B)$ und $(x \in B \Rightarrow x \in C)$ wahr sind, so gilt das auch für die Implikation $(x \in A \Rightarrow x \in C)$. Das beweist (ii).

zu c): Wann sind zwei Mengen gleich? Antwort: $A = B$ gilt genau dann, wenn für alle x die Aussage $(x \in A \Leftrightarrow x \in B)$ wahr ist. Hierbei bedeutet " \Leftrightarrow " die zweiseitige Implikation, also "dann und nur dann, wenn". Nach Definition 3.3 haben wir somit

$$A = B \Leftrightarrow (A \subset B \text{ und } B \subset A).$$

Insbesondere haben wir c) bewiesen. □

Definition 3.5 Es gibt eine Menge ohne Elemente. Sie wird mit \emptyset bezeichnet und heißt die **leere Menge**. Sie ist durch folgende Aussage charakterisiert:

$$\forall x \text{ gilt : } x \notin \emptyset.$$

Lemma 3.6 Ist M eine Menge, so gilt $\emptyset \subset M$.

Beweis: Da für alle Objekte x die Aussage $x \in \emptyset$ falsch ist, ist die Implikation $(x \in \emptyset \Rightarrow x \in M)$ stets wahr (ex falso quod libet). Also gilt $\emptyset \subset M$. □

Definition 3.7 Es sei I eine Menge (Indexmenge), und für jeden Index $i \in I$ sei eine Menge X_i gegeben. Dann kann man durch folgende Festsetzung eine neue Menge X bilden. Für alle x gelte

$$x \in X :\Leftrightarrow \exists i \in I : x \in X_i.$$

X heißt die **Vereinigung** der Mengen $X_i, i \in I$ und wird mit $\bigcup_{i \in I} X_i$ bezeichnet. Aus der Definition ergibt sich, dass $X = \emptyset$ gilt, wenn $I = \emptyset$ ist. Ist $n \in \mathbb{N}, n > 0$ und $I = \{1, \dots, n\}$, so schreibt man auch

$$X_1 \cup \dots \cup X_n = \bigcup_{i=1}^n X_i = \bigcup_{i \in I} X_i.$$

Insbesondere ist die Vereinigung zweier Mengen A, B definiert durch

$$A \cup B = \{x \mid x \in A \text{ oder } x \in B\}.$$

Hierbei ist “oder” das nicht ausschließende Oder. Die Wahrheitstafel für “oder” hat also die folgende Form:

p	q	p oder q
1	1	1
0	1	1
1	0	1
0	0	0

(1 = wahr, 0 = falsch)

Definition 3.8 Es sei I eine nichtleere Indexmenge, und für jedes $i \in I$ sei eine Menge X_i gegeben. Der **Durchschnitt**

$$X = \bigcap_{i \in I} X_i$$

der Mengen $X_i, i \in I$ wird durch

$$x \in X \Leftrightarrow \forall i \in I : x \in X_i$$

definiert. Für $I = \{1, \dots, n\}$ ($n > 0$) schreibt man auch

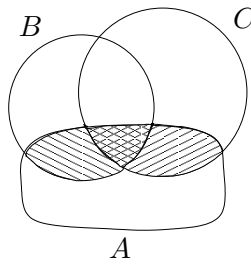
$$X_1 \cap \dots \cap X_n = \bigcap_{i=1}^n X_i = \bigcap_{i \in I} X_i.$$

Für zwei Mengen A, B ist

$$A \cap B = \{x \mid x \in A \text{ und } x \in B\}.$$

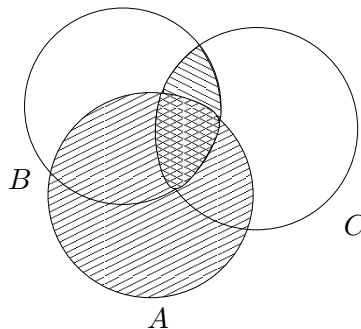
A und B heißen **disjunkt**, wenn $A \cap B = \emptyset$ gilt.

Als Übung überlege man sich die Distributivgesetze



$$A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$$

und



$$A \cup (B \cap C) = (A \cup B) \cap (A \cup C).$$

Bemerkung 3.9 Ist M eine Menge, $A(x)$ eine Aussage über x , so kann man die Teilmenge $N \subset M$ bilden, die aus allen Elementen $x \in M$ besteht, für die die Aussage $A(x)$ wahr ist, also:

$$x \in N :\Leftrightarrow x \in M \text{ und } A(x).$$

In Zeichen schreibt man dann

$$N = \{x \in M | A(x)\}.$$

Beispiel 3.10 a) $\{n \in \mathbb{Z} | \exists m \in \mathbb{Z} : n = 2m\}$ ist die Menge der geraden ganzen Zahlen,

b) $\{n \in \mathbb{Z} | n^2 = 1\} = \{-1, 1\},$

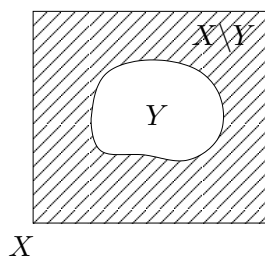
c) $\{n \in \mathbb{Z} | n \geq 0\} = \mathbb{N}.$

Definition 3.11 Sind X, Y Mengen, so heißt $X \setminus Y = \{x \in X | x \notin Y\}$ die **Differenzmenge** (lies: “ X ohne Y ”).

Es gelten einfache Regeln für die mengentheoretischen Verknüpfungen. Wir geben nur einige Beispiele: Seien A, B , Teilmengen einer Menge X . Dann gilt

$$\begin{aligned} X \setminus (A \cup B) &= (X \setminus A) \cap (X \setminus B), \\ X \setminus (A \cap B) &= (X \setminus A) \cup (X \setminus B), \\ (X \setminus A) \cup A &= X, \quad (X \setminus A) \cap A = \emptyset, \\ X \setminus (X \setminus A) &= A, \text{ usw.} \end{aligned}$$

Von fundamentaler Bedeutung ist der Begriff des geordneten Paares (von zwei Objekten), welcher die Definition des kartesischen Produktes zweier Mengen ermöglicht.



Definition 3.12

a) Seien x, y irgendwelche Objekte. Aus x und y kann man ein neues Element bilden, das **geordnete Paar** (auch kurz: **Paar**)

$$(x, y).$$

Dies bekommt einen Sinn, wenn man definiert, wann zwei Paare gleich sind. Seien dazu x', y' zwei weitere Objekte. Dann wird vereinbart

$$(x, y) = (x', y') :\Leftrightarrow x = x' \text{ und } y = y'.$$

Wir bemerken sogleich, dass (x, y) und (y, x) verschiedene Objekte sind, wenn x und y verschieden sind. Es gilt: $(x, y) = (y, x) \Leftrightarrow x = y$.

Das **kartesische Produkt** $X \times Y$ zweier Mengen X und Y wird nun definiert als

$$X \times Y = \{(x, y) | x \in X \text{ und } y \in Y\}.$$

b) Analog wird das **kartesische Produkt**

$$X_1 \times \dots \times X_n = \{(x_1, \dots, x_n) | x_i \in X_i \text{ für } i = 1, \dots, n\}$$

der Mengen X_1, \dots, X_n ($n \in \mathbb{N}, n \geq 1$) definiert. Hierbei wird vereinbart

$$(x_1, \dots, x_n) = (x'_1, \dots, x'_n) \Leftrightarrow \forall i = 1, \dots, n : x_i = x'_i.$$

Die Elemente von $X_1 \times \dots \times X_n$ heißen **n -Tupel** (genauer: **geordnete n -Tupel**). Ist $X = X_1 = \dots = X_n$, so ist die Schreibweise

$$X^n = \underbrace{X \times \dots \times X}_{n\text{-mal}} = X_1 \times \dots \times X_n$$

üblich.

Vertraut ist uns die Zahlenebene $\mathbb{R}^2 = \mathbb{R} \times \mathbb{R}$ und auch der n -dimensionale Zahlenraum \mathbb{R}^n .

Grundlegend für die moderne Mathematik ist der Abbildungsbegriff.

Definition 3.13 Seien X, Y Mengen. Eine **Abbildung** f von X nach Y (in Zeichen: $f : X \rightarrow Y$) ist eine Vorschrift, die jedem Element $x \in X$ genau ein Element $y = f(x) \in Y$ zuordnet.

Das Element $f(x)$ heißt das **Bild** von x unter f ; man schreibt $x \mapsto f(x)$.

X heißt der **Definitionsbereich** von f und Y heißt der **Wertebereich** von f .

$$G_f = \{(x, y) \in X \times Y | y = f(x)\} \subset X \times Y$$

heißt der **Graf** von f . Eine Teilmenge $M \subset X \times Y$ ist genau dann der Graf einer Abbildung von X nach Y , wenn gilt: Zu jedem $x \in X$ gibt es genau ein $y \in Y$, so dass $(x, y) \in M$.

Zwei Abbildungen $f : X \rightarrow Y, g : X' \rightarrow Y'$ sind genau dann gleich, wenn gilt:

- (i) $X = X'$,
- (ii) $Y = Y'$ und
- (iii) Für alle $x \in X$ gilt: $f(x) = g(x)$.

Achtung:

$$\begin{aligned} f & : \mathbb{R} \rightarrow \{0\}, f(x) := 0 \quad \text{und} \\ g & : \mathbb{R} \rightarrow \mathbb{R}, g(x) = 0 \end{aligned}$$

sind verschiedene Abbildungen, weil ihre Wertebereiche verschieden sind.

Mit $\text{Abb}(X, Y)$ wird die Menge aller Abbildungen $f : X \rightarrow Y$ bezeichnet. Wenn wir $f \in \text{Abb}(X, Y)$ sagen, meinen wir also: $f : X \rightarrow Y$ ist eine Abbildung von X nach

Y . Statt " $f : X \rightarrow Y$ " ist auch die Schreibweise $X \xrightarrow{f} Y$ gebräuchlich. Offensichtlich gilt für Elemente $f, g \in \text{Abb}(X, Y)$:

$$f = g \Leftrightarrow G_f = G_g.$$

Definition 3.14 Es sei $f : X \rightarrow Y$ eine Abbildung von X nach Y . $M \subset X$, $N \subset Y$ seien Teilmengen. Dann heißt

$$f(M) = \{y \in Y \mid \exists x \in M : y = f(x)\} = \{f(x) \mid x \in M\}$$

die **Bildmenge** von M unter f und

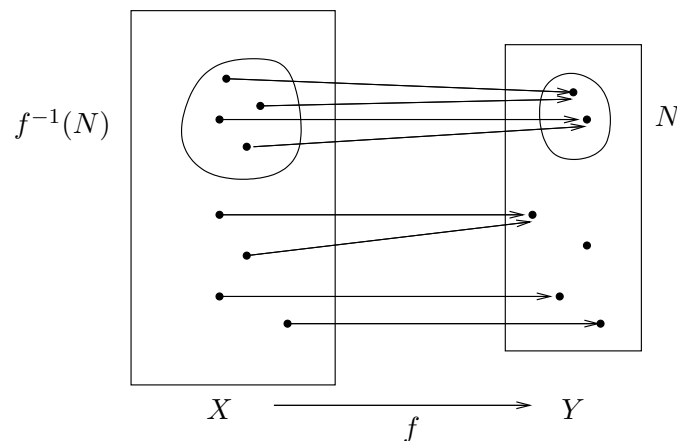
$$f^{-1}(N) = \{x \in X \mid f(x) \in N\}$$

die **Urbildmenge** von N unter f .

$$f(X) = \{f(x) \mid x \in X\} = \{y \in Y \mid f^{-1}(\{y\}) \neq \emptyset\}$$

nennt man das **Bild von f** (= Bildmenge von X unter f) und schreibt dafür bisweilen $\text{Im}f$.

Die Urbildmenge $f^{-1}(\{y\})$ einer einelementigen Teilmenge $\{y\} \subset Y$ nennt man auch die **Faser von f über y** . Es gilt: $f^{-1}(\{y\}) \neq \emptyset \Leftrightarrow y \in \text{Im}f$.



Beispiel 3.15 a) Es sei $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ die Abbildung mit der Vorschrift $f(x, y) = 2x + y$. Für $c \in \mathbb{R}$ ist dann $f^{-1}(\{c\}) = \{(x, y) \in \mathbb{R}^2 \mid 2x + y = c\}$ eine Gerade in \mathbb{R}^2 . Ist $d \in \mathbb{R}$, $d \neq c$, so ist

$$f^{-1}(\{c\}) \cap f^{-1}(\{d\}) = \emptyset.$$

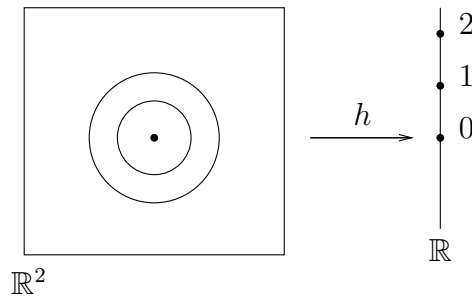
Also bilden die Fasern der Abbildung $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ eine Schar paralleler Geraden.

$$\begin{array}{ccc} & & c \\ & 0 & f \\ & & 0 \\ & & d \\ & & \mathbb{R} \\ \mathbb{R}^2 & & \end{array}$$

- b) $g : \mathbb{R}^2 \rightarrow \mathbb{R}$, $g(x, y) = x^2 - y^2 = (x - y)(x + y)$ hat als Faser über $0 \in \mathbb{R}$ eine Vereinigung von zwei Geraden:

$$\begin{aligned} g^{-1}(\{0\}) &= \{(x, y) \in \mathbb{R}^2 \mid x - y = 0 \text{ oder } x + y = 0\} \\ &= \{(x, y) \in \mathbb{R}^2 \mid x = y\} \cup \{(x, y) \in \mathbb{R}^2 \mid x = -y\}. \end{aligned}$$

Für $c \in \mathbb{R} \setminus \{0\}$ ist $g^{-1}(\{c\})$ eine Hyperbel (mit beiden Ästen).



- c) Schließlich betrachten wir noch

$$h : \mathbb{R}^2 \rightarrow \mathbb{R}, \quad h(x, y) = x^2 + y^2.$$

Das Bild von h ist $\mathbb{R}_{\geq 0} = \{x \in \mathbb{R} \mid x \geq 0\}$. Es gilt $h^{-1}(\{0\}) = \{(0, 0)\}$. Für $c \in \mathbb{R}$, $c > 0$ ist $h^{-1}(\{c\})$ der Kreis vom Radius \sqrt{c} um den Punkt $(0, 0)$.

Definition 3.16 Eine Abbildung $f : X \rightarrow Y$ heißt

- surjektiv** : $\Leftrightarrow f(X) = Y$, d.h.:
zu jedem Element $y \in Y$ gibt es mindestens ein Element $x \in X$ mit $f(x) = y$.
- injektiv** : $\Leftrightarrow \forall x, x' \in X : f(x) = f(x') \Rightarrow x = x'$
(oder anders formuliert: $\forall x, x' \in X : x \neq x' \Rightarrow f(x) \neq f(x')$).
- bijektiv** : $\Leftrightarrow f$ ist surjektiv und injektiv.

Definition 3.17 Es sei $f : X \rightarrow Y$ eine bijektive Abbildung. Dann gibt es zu jedem Element $y \in Y$ genau ein Element $x \in X$ mit $f(x) = y$. Dieses Element x wird mit $f^{-1}(y)$ bezeichnet. Damit ist eine Abbildung

$$f^{-1} : Y \rightarrow X$$

definiert. Es gilt: $f^{-1}(\{y\}) = \{f^{-1}(y)\}$, anders formuliert:

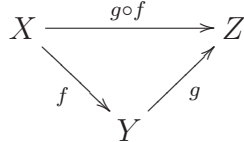
$$\forall x \in X, y \in Y : \quad f(x) = y \Leftrightarrow x = f^{-1}(y).$$

f^{-1} heißt die **Umkehrabbildung** von f .

Eine surjektive Abbildung $f : X \rightarrow Y$ ist genau dann bijektiv, wenn jede Faser von f aus genau einem Element besteht.

Definition 3.18

- a) Es seien $f : X \rightarrow Y, g : Y \rightarrow Z$ Abbildungen. Die **Komposition** $g \circ f : X \rightarrow Z$ von f und g wird folgendermaßen definiert:

$$(g \circ f)(x) := g(f(x)) \quad \text{für alle } x \in X;$$


ist ein kommutatives Diagramm von Abbildungen.

- b) Mit $id_X : X \rightarrow X$ wird die Abbildung $x \mapsto x$ bezeichnet. Sie heißt die **identische Abbildung** auf X . Ihr Graf ist die Diagonale $\Delta_X = \{(x, x) | x \in X\} \subset X \times X$.
- c) Sind X', Y' Mengen und $X \subset X', Y \subset Y'$ Teilmengen, $f : X \rightarrow Y, g : X' \rightarrow Y'$ Abbildungen, so heißt f **Einschränkung** von g und g **Fortsetzung** von f , wenn für alle $x \in X$ gilt:

$$f(x) = g(x).$$

Bezeichnet man mit $i : X \rightarrow X'$ und $j : Y \rightarrow Y'$ die Inklusionsabbildungen $i(x) = x, j(y) = y$, die sich von den identischen Abbildungen nur dadurch unterscheiden, dass Definitionsbereich und Wertebereich nicht notwendigerweise übereinstimmen, sondern der Definitionsbereich eine Teilmenge des Wertebereichs ist, so erhält man folgendes kommutatives Diagramm von Abbildungen

$$\begin{array}{ccc} X & \xrightarrow{f} & Y \\ i \downarrow & & \downarrow j \\ X' & \xrightarrow{g} & Y' \end{array} \quad \text{d.h., es gilt } j \circ f = g \circ i.$$

Lemma 3.19

- a) Die Komposition von Abbildungen ist eine assoziative Verknüpfung: Für Abbildungen $f : X \rightarrow Y, g : Y \rightarrow Z, h : Z \rightarrow W$ gilt

$$h \circ (g \circ f) = (h \circ g) \circ f.$$

- b) Für jede Abbildung $f : X \rightarrow Y$ gilt: $f \circ id_X = f, id_Y \circ f = f$.

Beweis:

- a) 1. W ist der Wertebereich von $h \circ (g \circ f)$ und von $(h \circ g) \circ f$.
2. Für alle $x \in X$ gilt:

$$\begin{aligned} (h \circ (g \circ f))(x) &= h((g \circ f)(x)) = h(g(f(x))) \\ &= (h \circ g)(f(x)) = ((h \circ g) \circ f)(x). \end{aligned}$$

Aus 1. und 2. folgt $h \circ (g \circ f) = (h \circ g) \circ f$.

- b) ist trivial. □

Lemma 3.20 Seien X, Y nichtleere Mengen, und $f : X \rightarrow Y$ sei eine Abbildung. Dann gilt:

- a) f ist injektiv $\Leftrightarrow \exists g : Y \rightarrow X$, so dass $g \circ f = id_X$
(d.h. f besitzt ein Linksinverses g).
- b) f ist surjektiv $\Leftrightarrow \exists g : Y \rightarrow X$, so dass $f \circ g = id_Y$
(d.h. f besitzt ein Rechtsinverses g).
- c) f ist bijektiv $\Leftrightarrow \exists g : Y \rightarrow X$, so dass $g \circ f = id_X$ und $f \circ g = id_Y$.
(d.h. f besitzt ein zweiseitiges Inverses g).

Beweis: zu a): Sei f injektiv; $g : Y \rightarrow X$ werde folgendermaßen definiert

$$g(y) = \begin{cases} x, & \text{falls } y \in f(X) \text{ und } x \text{ das Urbild von } y \text{ ist,} \\ \text{beliebig,} & \text{falls } y \in Y \setminus f(X). \end{cases}$$

Dann gilt offensichtlich $g \circ f = id_X$.

Sei nun umgekehrt $g \circ f = id_X$ für eine Abbildung $g : Y \rightarrow X$. Für $x, x' \in X$ gilt dann: Ist $f(x) = f(x')$, so ist $g(f(x)) = g(f(x'))$. Wegen $g \circ f = id_X$ folgt also $x = x'$. Die Abbildung f ist daher injektiv.

zu b): Es sei f surjektiv. Dann kann man zu jedem $y \in Y$ ein $x \in X$ mit $f(x) = y$ wählen. Dieses Element x nenne man $g(y)$. Dann hat man eine Abbildung $g : Y \rightarrow X$ mit $f \circ g = id_Y$ (Natürlich ist g nicht eindeutig bestimmt, wenn f mindestens eine Faser mit mehr als nur einem Element besitzt.).

Sei nun umgekehrt $f \circ g = id_Y$ für eine Abbildung $g : Y \rightarrow X$. Dann ist f offensichtlich surjektiv, denn für $y \in Y$ ist $x := g(y) \in X$ mit $f(x) = f(g(y)) = id_Y(y) = y$.

zu c): Sei f bijektiv und f^{-1} die Umkehrabbildung von f . Dann gilt

$$f(x) = y \quad \Leftrightarrow \quad x = f^{-1}(y).$$

Also gilt $f^{-1}(f(x)) = x$ und $f(f^{-1}(y)) = y$. f^{-1} ist ein zweiseitiges Inverses von f . Sei nun f eine Abbildung, die ein zweiseitiges Inverses $g : Y \rightarrow X$ besitzt. Nach a) und b) ist f dann surjektiv und injektiv, also bijektiv. \square

Bemerkung 3.21 Es gibt höchstens ein zweiseitiges Inverses von f . Denn: Sind $g, g' : Y \rightarrow X$ zwei zweiseitige Inverse von f , so gilt

$$g' = id_X \circ g' = (g \circ f) \circ g' = g \circ (f \circ g') = g \circ id_Y = g.$$

Korollar 3.22

- a) Ist $f : X \rightarrow Y$ bijektiv, so ist auch f^{-1} bijektiv, und es gilt $(f^{-1})^{-1} = f$.
- b) Sind $f : X \rightarrow Y, g : Y \rightarrow Z$ bijektiv, so ist auch $g \circ f : X \rightarrow Z$ bijektiv, und es gilt

$$(g \circ f)^{-1} = f^{-1} \circ g^{-1}.$$

Beweis: Wir beweisen nur b).

Wir müssen nur zeigen, dass $f^{-1} \circ g^{-1}$ ein zweiseitiges Inverses von $g \circ f$ ist. Das ist eine einfache Rechnung

$$(f^{-1} \circ g^{-1}) \circ (g \circ f) = f^{-1} \circ (g^{-1} \circ g) \circ f = f^{-1} \circ id_Y \circ f = f^{-1} \circ f = id_X$$

und

$$(g \circ f) \circ (f^{-1} \circ g^{-1}) = g \circ (f \circ f^{-1}) \circ g^{-1} = g \circ id_X \circ g^{-1} = g \circ g^{-1} = id_Z.$$

□

Wir schließen diesen Abschnitt mit einigen Bemerkungen über endliche Mengen.

Definition 3.23 Eine Menge M heißt **endlich** genau dann, wenn es eine natürliche Zahl $n \in \mathbb{N}$ und eine bijektive Abbildung $\{1, \dots, n\} \rightarrow M$ gibt (dabei sei $\{1, \dots, n\} = \emptyset$, wenn $n = 0$ ist). n heißt dann die Anzahl der Elemente von M , in Zeichen: $n = \text{card}M = |M|$. Ist M nicht endlich, so heißt M **unendlich**.

Offensichtlich gilt

Satz 3.24 Es sei M eine Menge. Folgende Aussagen sind äquivalent:

- (1) M ist endlich.
- (2) Jede injektive Abbildung $f : M \rightarrow M$ ist auch surjektiv.
- (3) Jede surjektive Abbildung $f : M \rightarrow M$ ist auch injektiv.

Bemerkung 3.25 Eine Menge M ist also unendlich, wenn man M bijektiv auf eine echte Teilmenge $N \subset M$ abbilden kann.

Muss man Satz 3.24 beweisen? Es ist doch klar, dass eine injektive Abbildung $\{1, \dots, n\} \rightarrow \{1, \dots, n\}$ automatisch surjektiv ist. Oder? Wir wollen dennoch einen Beweis präsentieren:

Zunächst beweisen wir (1) \Rightarrow (2): Es sei also M eine endliche Menge und $f : M \rightarrow M$ sei injektiv. $y \in M$ sei beliebig. Wir suchen ein $x \in M$ mit $f(x) = y$. Wir wollen also die Gleichung $f(x) = y$ in der Unbekannten x lösen. Es geht hier um einen Existenzbeweis. Unsere Strategie ist auch oft erfolgreich in komplizierteren Situationen anwendbar. Das Stichwort heißt *Iteration*. Vielleicht ist ihnen das Newtonsche Iterationsverfahren aus der Schule bekannt. In unserer Situation ist alles sehr einfach: Sei $f^0 = id_M$ und für $k > 0$ sei

$$f^k = f \circ f^{k-1}.$$

Wir betrachten die Folge $f^0(y), f(y), f^2(y), f^3(y), \dots$, und hoffen in dieser Folge eine Lösung unseres Problems zu finden. Da M endlich ist, gibt es zwei verschiedene Zahlen $l > k \geq 1$ mit

$$f^l(y) = f^k(y).$$

(Andernfalls wäre

$$\mathbb{N} \rightarrow M, \quad k \mapsto f^k(y)$$

eine injektive Abbildung, was der Endlichkeit von M widerspricht.) Aus $f(f^{l-1}(y)) = f^l(y) = f^k(y) = f(f^{k-1}(y))$ folgt $f^{l-1}(y) = f^{k-1}(y)$, weil f injektiv ist. Ist $k-1 \geq 1$, so fährt man fort und erhält schließlich $f^{l-k}(y) = y$, also $f(x) = y$ mit $x := f^{l-k-1}(y)$.

(2) \Rightarrow (3): Sei $f : M \rightarrow M$ surjektiv. Nach 3.20 b) gibt es ein Rechtsinverses $g : M \rightarrow M : f \circ g = id_M$. g ist dann injektiv, nach (2) also bijektiv, also ist $f = f \circ g \circ g^{-1} = g^{-1}$ auch bijektiv.

(3) \Rightarrow (2): beweist man genauso.

(2) \Rightarrow (1): Es sei M unendlich. Dann gibt es eine injektive Abbildung $h : \mathbb{N} \rightarrow M$. Es sei $N = h(\mathbb{N})$. Die Einschränkung $H : \mathbb{N} \rightarrow N$ mit $h(x) = H(x)$ ist dann bijektiv. Nun werde

$$f : M \rightarrow M$$

definiert durch

$$f(x) = \begin{cases} x, & \text{falls } x \in M \setminus N; \\ h(2n), & \text{falls } x = h(n), n \in \mathbb{N}. \end{cases}$$

Das Diagramm

$$\begin{array}{ccc} M & \xrightarrow{f} & M \\ \uparrow h & & \uparrow h \\ \mathbb{N} & \xrightarrow{(n \mapsto 2n)} & \mathbb{N} \end{array} \quad \text{ist kommutativ.}$$

Offensichtlich ist f injektiv, aber nicht surjektiv, denn $h(2n+1) \notin f(M)$. Damit haben wir die Implikation (nicht (1) \Rightarrow nicht (2)) bewiesen, was natürlich äquivalent zu ((2) \Rightarrow (1)) ist. \square

Definition 3.26

a) Eine unendliche Menge M heißt **abzählbar** unendlich \Leftrightarrow
 \exists bijektive Abbildung $f : \mathbb{N} \rightarrow M$.

b) Eine unendliche Menge M , die nicht abzählbar ist, heißt **überabzählbar**.

Beispiel 3.27 \mathbb{Q} ist abzählbar, aber \mathbb{R} ist überabzählbar (siehe [10]).

Übungen

1. Sei $A = \{+, -\}$, $\mathbb{N}_+ = \{n \in \mathbb{N} \mid n > 0\}$ und $B = \bigcup_{n \in \mathbb{N}_+} A^n$ die Menge der endlichen Folgen (n -Tupel, $n > 0$) von Plus- und Minuszeichen. $g : A \rightarrow A$ sei die Abbildung $g(+)= -, g(-)= +$ und $f : B \rightarrow B$ werde definiert durch

$$f(a_1, \dots, a_n) = (a_1, g(a_1), a_2, g(a_2), \dots, a_n, g(a_n))$$

für alle $a_1, \dots, a_n \in A$. Es sei $f^2 = f \circ f$. Zeigen Sie: Ist $f^2(a_1, \dots, a_n) = (b_1, \dots, b_{4n})$, so kommt in jedem Abschnitt $(b_i, b_{i+1}, b_{i+2}, b_{i+3}, b_{i+4})$ ($1 \leq i \leq 4(n-1)$) das Paar $(+, +)$ oder $(-, -)$ vor.

2. Besitzt $f : \mathbb{N} \rightarrow \mathbb{N}$, $f(n) = n^2$ ein Rechtsinverses? Geben Sie zwei Linksinverse von f an. Ist die Menge $\{g : \mathbb{N} \rightarrow \mathbb{N} \mid g \circ f = id_{\mathbb{N}}\}$ endlich oder unendlich?
3. Man überlege sich, dass für jede Menge M

$$\text{Abb}(\emptyset, M) = \{\emptyset\} \neq \emptyset,$$

aber für jede nichtleere Menge M

$$\text{Abb}(M, \emptyset) = \emptyset$$

gilt.

4. Seien A, B, C Mengen. Zeigen Sie

- a) $A \cap (B \setminus C) = (A \cap B) \setminus (A \cap C)$,
 b) $(A \cap B) \cup C = A \cap (B \cup C) \Leftrightarrow C \subset A$.

5. Seien X, Y, Z Mengen. Konstruieren Sie eine bijektive Abbildung

$$\Phi : \text{Abb}(X \times Y, Z) \rightarrow \text{Abb}(Y, \text{Abb}(X, Z)).$$

6. Es sei $f : X \rightarrow Y$ eine Abbildung, I eine Indexmenge und für $i \in I$ sei X_i eine Teilmenge von X und Y_i eine Teilmenge von Y . Beweisen Sie:

- a) $f\left(\bigcup_{i \in I} X_i\right) = \bigcup_{i \in I} f(X_i)$,
 b) $f\left(\bigcap_{i \in I} X_i\right) \subset \bigcap_{i \in I} f(X_i)$,
 c) $f^{-1}\left(\bigcup_{i \in I} Y_i\right) = \bigcup_{i \in I} f^{-1}(Y_i)$,
 d) $f^{-1}\left(\bigcap_{i \in I} Y_i\right) = \bigcap_{i \in I} f^{-1}(Y_i)$.

Zeigen Sie, dass in b) im allgemeinen nicht das Gleichheitszeichen stehen darf.

- e) Für $A \subset X$ gilt: $A \subset f^{-1}(f(A))$. Ist f injektiv, so gilt stets $A = f^{-1}(f(A))$.
 f) Für $B \subset Y$ gilt: $f(f^{-1}(B)) \subset B$. Ist $B \subset f(X)$, so gilt: $f(f^{-1}(B)) = B$.

7. Es sei $f : X \rightarrow Y$ eine Abbildung. Für $b \in f(X)$ sei $X_b = f^{-1}(\{b\})$. Beweisen Sie:

- (i) $X = \bigcup_{b \in f(X)} X_b$,

(ii) $b, b' \in f(X)$, $b \neq b' \Rightarrow X_b \cap X_{b'} = \emptyset$.

(Die Fasern der Abbildung f bilden eine Zerlegung von X .)

8. Beweisen Sie:

- a) Sind X und Y abzählbar unendlich, so ist auch $X \times Y$ abzählbar unendlich.
- b) \mathbb{Q} ist abzählbar unendlich.
- c) Die Menge der reellen algebraischen Zahlen $\{a \in \mathbb{R} \mid \exists n \in \mathbb{N}_+ \exists a_0, \dots, a_{n-1} \in \mathbb{Q}, \text{ so dass } a^n + a_{n-1}a^{n-1} + \dots + a_1a + a_0 = 0\}$ ist abzählbar unendlich.

9. a) Bestimmen Sie die Fasern der Abbildung

$$f : \mathbb{R}^2 \rightarrow \mathbb{R}^2, f(x, y) = (x^2 + y^2, xy).$$

- b) Untersuche die Abbildung $f : \mathbb{R} \rightarrow \mathbb{R}$, $f(x) = x^3 - 2x$ und ihre Einschränkung $g : \mathbb{Z} \rightarrow \mathbb{Z}$, $g(x) = x^3 - 2x$ auf Injektivität, Surjektivität, Bijektivität.

10. Seien X, Y endliche Mengen. Bestimmen Sie

$$|X \times Y|, \quad |\text{Abb}(X, Y)|, \quad |\mathcal{P}(X)| \quad |S(X)|,$$

wobei $\mathcal{P}(X) = \{A \mid A \subset X\}$ und $S(X) = \{f \in \text{Abb}(X, X) \mid f \text{ ist bijektiv}\}$.

4 Vollständige Induktion

Die Menge \mathbb{N} der natürlichen Zahlen ist linear geordnet:

$$0 < 1 < 2 < 3 < \dots < n < n + 1 < \dots$$

Zwischen einer Zahl $n \in \mathbb{N}$ und ihrem Nachfolger $n + 1$ gibt es keine weiteren natürlichen Zahlen. Unmittelbar einleuchtend ist daher der so genannte Satz vom kleinsten Element.

Satz 4.1 (Satz von kleinsten Element, Minimumsprinzip): Jede nichtleere Teilmenge $N \subset \mathbb{N}$ besitzt ein kleinstes Element. Das soll heißen: Es gibt ein $m \in N$, so dass $m \leq n$ für alle $n \in N$.

Beachte:

Für $\mathbb{Q}_{\geq 0} = \{x \in \mathbb{Q} \mid x \geq 0\}$ gilt das Minimumsprinzip nicht; z.B. besitzt $N = \{x \in \mathbb{Q}_{\geq 0} \mid x > 0\}$ kein kleinstes Element. 0 gehört ja nicht zu N und einen Nachfolger von 0 gibt es in $\mathbb{Q}_{\geq 0}$ nicht! Immerhin hat N mit 0 eine größte untere Schranke in $\mathbb{Q}_{\geq 0}$. Betrachtet man aber z.B. $M = \{x \in \mathbb{Q}_{\geq 0} \mid x^2 \geq 2\}$, so besitzt M nicht einmal eine größte untere Schranke in $\mathbb{Q}_{\geq 0}$ (denn $\sqrt{2}$ ist nicht rational).

Aus dem Minimumsprinzip kann man das Prinzip der vollständigen Induktion, das eine wichtige Beweismethode darstellt, ableiten.

Satz 4.2 (Prinzip der vollständigen Induktion): Für jede natürliche Zahl n sei eine Aussage $A(n)$ formuliert. Folgende Bedingungen seien erfüllt:

- (1) Die Aussage $A(0)$ ist wahr.
- (2) Für jede natürliche Zahl $n \in \mathbb{N}$ gilt:
Ist die Aussage $A(n)$ wahr, so ist auch $A(n + 1)$ wahr.

Dann gilt:

Die Aussage $A(n)$ ist für alle $n \in \mathbb{N}$ gültig.

Beweis: Wir führen einen Widerspruchsbeweis. Sei $N = \{n \in \mathbb{N} \mid A(n) \text{ ist falsch}\}$. Nehmen wir nun an, dass $A(n)$ nicht für alle $n \in \mathbb{N}$ gültig ist. Dann ist also $N \neq \emptyset$. Nach dem Minimumsprinzip gibt es ein $m \in N$ mit $m \leq n$ für alle $n \in N$. Da nach (1) die Aussage $A(0)$ gültig ist, ist $0 \notin N$, also $0 < m$ und somit ist $m - 1 \in \mathbb{N}$ und wegen $m - 1 < m$ gilt $m - 1 \notin N$, d.h. die Aussage $A(m - 1)$ ist gültig; nach (2) ist dann auch $A(m)$ gültig (denn: $m = (m - 1) + 1$). Das steht aber im Widerspruch zu $m \in N$. Die Annahme $N \neq \emptyset$ ist also falsch; es gilt vielmehr $N = \emptyset$, d.h. $\forall n \in \mathbb{N}$: $A(n)$ ist wahr. \square

Eine kleine Abwandlung ist:

Um eine Aussage $A(n)$ für alle $n \geq n_0$ zu beweisen, zeigt man zunächst

- a) $A(n_0)$ gilt (Dies ist der **Induktionsanfang**.)
und dann:

- b) Für $n > n_0$ gilt: Ist $A(n-1)$ gültig, so ist auch $A(n)$ gültig (Dies ist der **Induktionsschluß** von $n-1$ auf n).

Oftmals wird auch folgende Form des Induktionsprinzips benutzt:

Satz 4.3 (Prinzip der vollständigen Induktion, zweite Form):

Für jede natürliche Zahl $n \in \mathbb{N}$ sei eine Aussage $A(n)$ gegeben. Die Aussage, $A(0)$ sei wahr und, man habe ein Verfahren, mit dem man für jedes $n > 0$ zeigen kann, dass $A(n)$ wahr ist, wenn $A(m)$ für alle $m \in \mathbb{N}$ mit $m < n$ wahr ist. Dann ist $A(n)$ für alle $n \in \mathbb{N}$ wahr.

Beweis: Falls $N = \{n \in \mathbb{N} | A(n) \text{ ist falsch}\}$ nichtleer ist, gibt es ein kleinstes Element $n_0 \in N$, $n_0 > 0$, und für alle $m \in \mathbb{N}$ mit $m < n_0$ ist somit $A(m)$ wahr; aber dann muss auch $A(n_0)$ wahr sein, Widerspruch! \square

Bemerkung: Für einen Induktionsbeweis ist es wichtig, den Induktionsanfang zu beweisen. Wir illustrieren dies an folgendem "Beweis":

Behauptung: Alle Studenten sind gleich groß.

"Beweis":

a) (Induktionsanfang) Für die leere Menge von Studenten und für eine Menge, die nur einen Studenten als Element enthält, ist die Behauptung wahr.

b) Jetzt führen wir den Induktionsschluß durch:

Wir setzen also voraus: Für eine Menge von n Studenten ist die Behauptung richtig, d.h. in einer Menge mit n Studenten sind alle Studenten gleich groß.

Wir zeigen: Dann ist die Behauptung auch für eine Menge mit $n+1$ Studenten richtig. Das geht so: Die ersten n Studenten und die letzten n Studenten haben jeweils dieselbe Größe. Da der zweite Student zu beiden Mengen gehört, haben also alle $n+1$ Studenten dieselbe Größe.

Was ist hier falsch?

Antwort: Damit man obigen Induktionsschluss durchführen kann, muss $n > 2$ gelten, man muss also als Induktionsanfang $n = 2$ wählen!

Natürlich reicht es auch nicht aus, eine Aussage $A(n)$ für alle $n \in \mathbb{N}$ dadurch zu beweisen, dass man den Beweis für eine endliche, und sei es noch so große, Anzahl von Fällen durchführt. Das wäre sozusagen eine *unvollständige* Induktion. Wir geben zwei Beispiele:

- a) $F_n = 2^{2^n} + 1$, $n \in \mathbb{N}$, heißt die n -te **Fermatsche Zahl**. $F_0 = 3$, $F_1 = 5$, $F_2 = 17$, $F_3 = 257$, $F_4 = 65537$ sind Primzahlen. Also vermutete Fermat: $\forall n \in \mathbb{N}$: F_n ist Primzahl. Aber der geniale Mathematiker Leonhard Euler (1707-1783) fand heraus, dass 641 ein Teiler von F_5 ist. Die Lösung sieht ganz einfach aus, sie zu finden war ein kleiner Geniestreich: $641 = 625 + 16 = 5^4 + 2^4$, also folgt $F_5 - 1 = 2^{32} = 2^4 2^{28} = (641 - 5^4) 2^{28} = 641 \cdot 2^{28} - (5 \cdot 2^7)^4 = 641 \cdot 2^{28} - 640^4 = 641 \cdot 2^{28} - (641 - 1)^4 = 641(2^{28} - 641^3 + 4 \cdot 641^2 - 6 \cdot 641 + 4) - 1$. Also ist 641 ein Teiler von F_5 . Es ist übrigens keine Zahl $n \geq 5$ bekannt, für die F_n eine Primzahl ist. Die unvorstellbar große Zahl $2^{2^{1945}} + 1$ hat den Teiler $5 \cdot 2^{1947} + 1$. Mit dem Computeralgebrasystem *Mathematica*[®] kann man

leicht die Primfaktorzerlegungen etwa von F_5, F_6, F_7, F_8 berechnen. Das war zu Eulers Zeiten natürlich nicht möglich.

- b) Die Zahl $n^2 - n + 41$ ist für $n = 1, 2, \dots, 40$ eine Primzahl, aber für $n = 41$ ergibt sich natürlich die Quadratzahl 41^2 .

Definition 4.4 a) Für $n \in \mathbb{N}$ sei die natürliche Zahl $n!$ (lies: n -Fakultät) definiert durch

$$0! = 1 \quad \text{und} \quad n! = 1 \cdot 2 \cdot 3 \cdot \dots \cdot n \quad \text{für } n \geq 1.$$

- b) Für $x \in \mathbb{R}$ und $k \in \mathbb{N}$ sei

$$\binom{x}{k} = \frac{x(x-1)\dots(x-k+1)}{k \cdot (k-1) \cdot \dots \cdot 1} \in \mathbb{R}$$

Ist $n \in \mathbb{N}$ und $n \geq k$, so ist

$$\binom{n}{k} = \frac{n!}{k!(n-k)!} \in \mathbb{N}. \quad (24)$$

$\binom{n}{k}$ heißt **Binomialkoeffizient**.

Offensichtlich gilt also $\binom{n}{k} = \binom{n}{n-k}$ für $0 \leq k \leq n$. Weiter gilt für $1 \leq k \leq n$ die Rekursionsformel

$$\binom{n}{k-1} + \binom{n}{k} = \binom{n+1}{k}. \quad (25)$$

Beweis:
$$\frac{n!}{(k-1)!(n-k+1)!} + \frac{n!}{k!(n-k)!} = \frac{k \cdot n! + (n-k+1)n!}{k!(n-k+1)!} = \binom{n+1}{k}. \quad \square$$

					1				
				1	1				
			1	2	1				
		1	3	3	1				
		1	4	6	4	1			
	1	5	10	10	5	1			
	1	6	15	20	15	6	1		
	1	6	15	20	15	6	1		

$$\binom{7}{0} \quad \binom{7}{1} \quad \binom{7}{2} \quad \binom{7}{3} \quad \binom{7}{4} \quad \binom{7}{5} \quad \binom{7}{6} \quad \binom{7}{7}$$

Das Pascalsche Dreieck

In der n -ten Zeile ($n = 0, 1, 2, \dots$) stehen die $n+1$ Binomialkoeffizienten $\binom{n}{0}, \binom{n}{1}, \dots, \binom{n}{n}$. Das Bildungsgesetz für das Dreieck ergibt sich aus der Rekursionsformel (25).

Als Übung zur vollständigen Induktion zeige man:

Für beliebige reelle Zahlen a, b und für $n \in \mathbb{N}$ gilt

$$(a+b)^n = a^n + \binom{n}{1}a^{n-1}b + \binom{n}{2}a^{n-2}b^2 + \dots + \binom{n}{k}a^{n-k}b^k + \dots + b^n. \quad (26)$$

Satz 4.5 Es sei X eine Menge mit k Elementen, Y eine Menge mit n Elementen, und es sei $k \leq n$. Dann gibt es $k! \binom{n}{k}$ injektive Abbildungen $f : X \rightarrow Y$.

Beweis: durch Induktion nach k :

- a) **Induktionsanfang:** Es sei $k = 0$. Dann ist $X = \emptyset$, und es gibt genau eine injektive Abbildung $\emptyset \rightarrow Y$. Da $0! \binom{n}{0} = 1$ gilt, ist der Satz für $k = 0$ bewiesen.
- b) **Induktionsschluß:** Es sei $k \in \mathbb{N}$ und, der Satz sei schon für k bewiesen. Wir zeigen, dass er dann auch für $k + 1$ gilt.

Dazu sei jetzt X eine Menge mit $k + 1$ Elementen und $k + 1 \leq n$. Da $k + 1 \geq 1$, ist $X \neq \emptyset$. Sei $a \in X$ fest gewählt. Dann ist $X' = X \setminus \{a\}$ eine Menge mit k Elementen, also gibt es nach Induktionsvoraussetzung $k! \binom{n}{k}$ injektive Abbildungen $F : X' \rightarrow Y$ und, da $|Y \setminus f(X')| = n - k$ ist, kann man jede solche Abbildung auf $n - k$ verschiedene Weisen zu einer injektiven Abbildung $g : X \rightarrow Y$ fortsetzen (Man kann $g(a) \in Y \setminus f(X')$ beliebig wählen.). Also gibt es $(n - k)k! \binom{n}{k}$ injektive Abbildungen von X nach Y . Man rechnet nun etwas:

$$(n - k)k! \binom{n}{k} = (n - k) \frac{n!}{(n - k)!} = \frac{n!}{(n - k - 1)!} = (k + 1)! \binom{n}{k + 1}.$$

Damit ist der Induktionsschluß durchgeführt und Satz 4.5 bewiesen. \square

Folgerung 4.6 Sei X eine Menge mit n Elementen. Dann gilt:

- a) $|S(X)| = n!$.
- b) Ist $0 \leq k \leq n$, so gibt es genau $\binom{n}{k}$ Teilmengen A von X mit $|A| = k$.
- c) $|\mathcal{P}(X)| = 2^n$.

Beweis: zu a): $S(X) = \{f : X \rightarrow X \mid f \text{ injektiv}\}$. Jetzt wende Satz 4.5 für $X = Y$ an.

zu b): Es sei $\mathcal{P}_k(X)$ die Menge aller Teilmengen $A \subset X$ mit k Elementen. $I_k(M)$ sei die Menge aller injektiven Abbildungen $f : \{1, \dots, k\} \rightarrow M$. Dann gilt:

$$I_k(X) = \bigcup_{A \in \mathcal{P}_k(X)} I_k(A)$$

ist disjunkte Vereinigung und

$$|I_k(A)| = k!, \quad |I_k(X)| = k! \binom{n}{k} \quad (\text{nach Satz 4.5}).$$

Da aber andererseits $|I_k(X)|$ die Summe der Anzahlen $|I_k(A)|$, $A \in \mathcal{P}_k(X)$ ist, folgt

$$|\mathcal{P}_k(X)| = |I_k(X)| / |I_k(A)| = \binom{n}{k}.$$

zu c):

$$|\mathcal{P}(X)| = |\mathcal{P}_0(X)| + \dots + |\mathcal{P}_n(X)| = \binom{n}{0} + \dots + \binom{n}{n} = (1 + 1)^n = 2^n.$$

\square

Übungen

1. Es sei $p_k(n) = 1^k + 2^k + 3^k + \dots + n^k$ die Summe der ersten n k -ten Potenzen, wobei $k = 1, 2, \dots$. Beweisen Sie durch Induktion über n :

a) $p_3(n) = p_1(n)^2$,

b) $p_5(n) = 2p_1(n)^4 - p_1(n)$.

2. Es sei M eine endliche Menge mit n Elementen. Beweisen Sie: Es gibt genau 3^n Paare (A, B) aus disjunkten Teilmengen $A, B \subset M$.
3. Es seien x_0, \dots, x_n Unbekannte und $k_0, \dots, k_n \in \mathbb{N}$. Dann heißt

$$x_0^{k_0} x_1^{k_1} \dots x_n^{k_n} \quad \text{ein Monom vom Grad } k = k_0 + \dots + k_n$$

in den Unbekannten x_0, \dots, x_n . Beweisen Sie: Es gibt $\binom{n+k}{k}$ verschiedene Monome vom Grad k in den Unbekannten x_0, \dots, x_n .

4. Beweisen Sie: Die Summe der ersten n ungeraden Zahlen ist eine Quadratzahl.

5 Äquivalenzrelationen

In diesem Abschnitt behandeln wir den wichtigen Begriff der **Äquivalenzrelation** und der **Quotientenbildung**.

Definition 5.1 Es sei X eine Menge. Eine **Relation** (genauer: **binäre Relation**) R auf X ist eine Vorschrift, die für jedes Paar $(x, y) \in X \times X$ festlegt, ob x bezüglich R in Relation zu y steht oder nicht. Ist x in Relation zu y , so schreiben wir kurz

$$xRy.$$

Das Symbol R wird hier als Relationszeichen verwendet.

Beispiel 5.2

a) Auf der Menge \mathbb{Z} haben wir z.B.

- i) die Gleichheitsrelation “=”.
- ii) die Kleiner-Relation “<”.
- iii) die Teiler-Relation “|”. Dabei wird definiert

$$a|b :\Leftrightarrow \exists c \in \mathbb{Z}, \text{ so daß } b = ac.$$

iv) die Kongruenzrelation “ $\equiv \text{ mod } 5$ ”. Dabei ist

$$a \equiv \text{ mod } 5 :\Leftrightarrow 5|a - b.$$

b) Eine Abbildung $f : X \rightarrow X$ definiert auch eine Relation “ \xrightarrow{f} ” auf X , nämlich

$$x \xrightarrow{f} y :\Leftrightarrow y = f(x).$$

Definition 5.3 Ist R eine Relation auf X , so heißt

$$G_R = \{(x, y) \in X \times X | xRy\}$$

der **Graph** von R .

Nach Definition ist die Abbildung

$$\begin{array}{ccc} \{\text{Relationen auf } X\} & \longrightarrow & \mathcal{P}(X \times X) \\ R & \longmapsto & G_R \end{array}$$

bijektiv. Die Aussagen xRy und $(x, y) \in G_R$ sind äquivalent. Nicht alle Relationen sind mathematisch wichtig oder interessant!

Aber von fundamentaler Bedeutung sind die Äquivalenzrelationen. Eine Äquivalenzrelation R auf einer Menge X ermöglicht es, von den speziellen Eigenschaften der Elemente von X zu abstrahieren und eine neue Menge Y einzuführen, in der zwischen verschiedenen Elementen der Menge X nicht mehr unterschieden wird, wenn sie bezüglich unserer Äquivalenzrelation R in Relation stehen.

Definition 5.4 Es sei X eine Menge. Eine Relation \sim auf X heißt **Äquivalenzrelation** \Leftrightarrow

- a) Für alle $x \in X$ gilt: $x \sim x$ (Reflexivität).
- b) Für alle $x, y \in X$ gilt: aus $x \sim y$ folgt $y \sim x$ (Symmetrie).
- c) Für alle $x, y, z \in X$ gilt: aus $x \sim y$ und $y \sim z$ folgt $x \sim z$ (Transitivität).

Gilt $x \sim y$, was äquivalent zu $y \sim x$ ist, so sagen wir x und y sind äquivalent (bezüglich \sim). Sind x und y nicht äquivalent, so schreiben wir $x \not\sim y$. Für $x \in X$ heißt

$$[x] = \{y \in X \mid x \sim y\}$$

die **Äquivalenzklasse** von x . Genauer schreibt man auch $[x]_{\sim}$ statt $[x]$ und sagt \sim -Äquivalenzklasse, statt Äquivalenzklasse.

Satz 5.5 Es sei \sim eine Äquivalenzrelation auf der Menge X . Dann gilt:

- a) Für $x, y \in X$ ist

$$[x] = [y], \quad \text{falls } x \sim y$$

$$\text{und } [x] \cap [y] = \emptyset, \quad \text{falls } x \not\sim y.$$

- b) X ist die Vereinigung der \sim -Äquivalenzklassen. In Zeichen,

$$X = \bigcup_{x \in X} [x].$$

a) und b) besagen, daß die \sim -Äquivalenzklassen eine Zerlegung von X in paarweise disjunkte Teilmengen bilden.

Beweis: zu a): Nach Definition ist die Aussage $x \sim y$ äquivalent zu $y \in [x]$ und wegen der Symmetrie auch zu $x \in [y]$. Um zu zeigen, daß aus $x \sim y$ die Gleichung

$$[x] = [y]$$

folgt, brauchen wir also nur zu zeigen:

Behauptung: Ist $y \in [x]$, so ist $[y] \subset [x]$.

Beweis: Sei $z \in [y]$, also $y \sim z$. Da auch $x \sim y$ gilt, impliziert die Transitivität $x \sim z$, d.h. $z \in [x]$. Damit ist $[y] \subset [x]$ bewiesen.

Jetzt zeigen wir, daß aus $[x] \cap [y] \neq \emptyset$ notwendig $x \sim y$ folgt. Sei nämlich $z \in [x] \cap [y]$. Dann ist $x \sim z$ und $y \sim z$. Wegen der Symmetrie ist auch $z \sim y$. Die Transitivität ergibt dann $x \sim y$.

Zu b): Wegen der Reflexivität ist $x \in [x]$ und, somit ist $X = \bigcup_{x \in X} [x]$. □

Definition 5.6 Es sei X eine Menge und \sim eine Äquivalenzrelation auf X . Mit X/\sim wird dann die Menge der \sim -Äquivalenzklassen bezeichnet:

$$X/\sim = \{[x] \mid x \in X\}.$$

X/\sim heißt auch die **Quotientenmenge** von X nach der Äquivalenzrelation \sim .

Die surjektive Abbildung $p : X \rightarrow X/\sim$ mit $x \mapsto p(x) = [x]$ heißt die **natürliche Abbildung** oder die **kanonische Projektion**.

Da nach Satz 2.3.5 $[x] = [y]$ zu $x \sim y$ äquivalent ist, sind die Fasern der Abbildung p genau die \sim -Äquivalenzklassen. Es gilt:

$$\begin{aligned} p^{-1}(\{[x]\}) &= \{y \in X \mid p(y) = [x]\} \\ &= \{y \in X \mid [y] = [x]\} \\ &= \{y \in X \mid y \sim x\} = [x]. \end{aligned}$$

Ist $y \in X/\sim$ und $x \in p^{-1}(\{y\})$, also $[x] = y$, so heißt x ein **Repräsentant** von y .

Ein **volles Repräsentantensystem** für die Äquivalenzrelation \sim ist eine Teilmenge $M \subset X$, so daß die Einschränkungabbildung

$$p|_M : M \rightarrow X/\sim$$

von p auf die Teilmenge M bijektiv ist. In der Regel gibt es viele verschiedene Repräsentantensysteme. Da p surjektiv ist, gibt es ein Rechtsinverses $q : X/\sim \rightarrow X$ (also $p \circ q = id_{X/\sim}$). Für solch ein q ist $q(X/\sim) \subset X$ ein volles Repräsentantensystem.

Beispiel 5.7

- a) Für $(x, y), (x', y') \in \mathbb{R}^2$ sei $(x, y) \sim (x', y')$, wenn $y = y'$ gilt. Dies ist eine Äquivalenzrelation und, die Äquivalenzklasse von $(a, b) \in \mathbb{R}^2$ ist $[(a, b)] = \{(x, y) \in \mathbb{R}^2 \mid y = b\}$, also eine Gerade. Ein volles Repräsentantensystem $M \subset \mathbb{R}^2$ ist z.B. eine Gerade $M = \{(x, y) \in \mathbb{R}^2 \mid x = cy\}$, wobei $c \in \mathbb{R}$.

Es gilt $M \cap [(a, b)] = \{(bc, b)\}$ für alle $(a, b) \in \mathbb{R}^2$. Also ist die Einschränkung $p|_M : M \rightarrow \mathbb{R}^2/\sim$ bijektiv.

- b) Für $x, y \in \mathbb{Z}$ sei $x \sim y$, wenn $x \equiv y \pmod{5}$ gilt. Für $x \in \mathbb{Z}$ ist dann

$$[x] = \{\dots x - 10, x - 5, x, x + 5, x + 10, \dots\}$$

und

$$\begin{aligned} M_1 &= \{0, 1, 2, 3, 4\}, \\ M_2 &= \{-2, -1, 0, 1, 2\}, \\ M_3 &= \{100, 51, 42, 103, -4\} \subset \mathbb{Z} \end{aligned}$$

sind Beispiele voller Repräsentantensysteme dieser Äquivalenzrelation. Natürlich gilt

$$\mathbb{Z}/\sim = \{[0], [1], [2], [3], [4]\}.$$

Dieses Beispiel wird uns noch oft begegnen.

Lemma 5.8 Es sei \sim eine Äquivalenzrelation auf X und $p : X \rightarrow X/\sim$ die kanonische Projektion. Eine Abbildung $f : X \rightarrow Y$ heißt \sim -invariant, wenn aus $x \sim x'$ stets $f(x) = f(x')$ folgt. Ist f eine \sim -invariante Abbildung, so gibt es genau eine Abbildung $\bar{f} : X/\sim \rightarrow Y$ mit $\bar{f}([x]) = f(x)$ für alle $x \in X$.

$$\begin{array}{ccc} X & \xrightarrow{f} & Y \\ & \searrow p & \nearrow \bar{f} \\ & X/\sim & \end{array}$$

ist kommutatives Diagramm, d.h. $\bar{f} \circ p = f$.

Beweis: Für $a \in X/\sim$ wähle man einen Repräsentanten $x \in X$; es gilt dann $a = [x]$. Wir setzen

$$\bar{f}(a) := f(x). \quad (*)$$

Dies hängt nicht von der Wahl von x ab. Ist nämlich x' ein anderer Repräsentant von a , also $a = [x']$, so gilt $x \sim x'$ und wegen der Invarianz von f , somit $f(x) = f(x')$. Man sagt: $\bar{f} : X/\sim \rightarrow Y$ ist durch (*) wohldefiniert. Nach Definition gilt

$$\bar{f} \circ p = f.$$

Damit ist die Existenz von \bar{f} bewiesen.

Um die Eindeutigkeit von \bar{f} zu beweisen, sei $h : X/\sim \rightarrow Y$ eine beliebige Abbildung mit

$$h \circ p = f.$$

Dann gilt $\bar{f} \circ p = h \circ p$, und da p surjektiv ist, folgt dann $\bar{f} = h$.

Das Lemma ist bewiesen. □

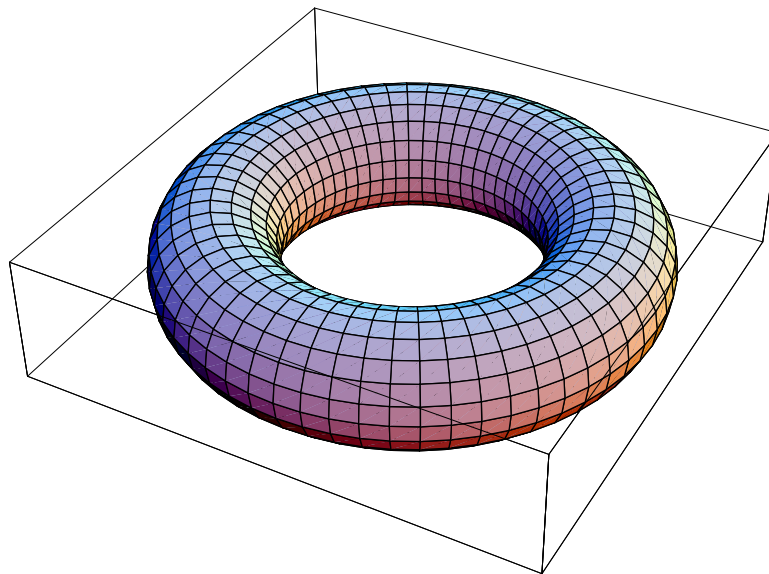
\bar{f} heißt die von f induzierte Abbildung.

Übungen

1. Es sei $f : X \rightarrow Y$ eine Abbildung. Für $x, x' \in X$ sei $x \sim x'$ definitionsgemäß genau dann, wenn $f(x) = f(x')$ gilt. Zeigen Sie: \sim ist eine Äquivalenzrelation auf X .

Sei $p : X \rightarrow X/\sim$ die kanonische Projektion. Zeigen Sie: f induziert eine injektive Abbildung $\bar{f} : X/\sim \rightarrow Y$ mit $\bar{f} \circ p = f$ und $\bar{f}(X/\sim) = f(X)$.

2. Für $(x, y), (x', y') \in \mathbb{R}^2$ sei $(x, y) \sim (x', y')$ genau dann, wenn $x - x'$ und $y - y'$ ganzzahlig sind. Zeigen Sie, daß \sim eine Äquivalenzrelation auf \mathbb{R}^2 ist und erklären Sie, wie man eine bijektive Abbildung von \mathbb{R}^2/\sim auf den **Torus** $T \subset \mathbb{R}^3$ konstruieren kann.



Dabei ist

$$T := \{(\cos \alpha \cos \beta + 3 \cos \alpha, \sin \alpha \cos \beta + 3 \sin \alpha, \sin \beta) \mid \alpha, \beta \in \mathbb{R}\}.$$

Zeigen Sie weiter, dass

$$M = \{(x, y) \in \mathbb{R}^2 \mid 0 \leq x < 1, 0 \leq y < 1\}$$

ein volles Repräsentantensystem von \sim ist.

3. Wieviele verschiedene Äquivalenzrelationen gibt es auf den Mengen $\{a, b\}$, $\{a, b, c\}$, $\{a, b, c, d\}$?
4. Zeigen Sie: Durch $(x, y) \sim (x', y') \Leftrightarrow xy = x'y'$ ist auf \mathbb{R}^2 eine Äquivalenzrelation gegeben, und $M = \{(x, y) \in \mathbb{R}^2 \mid y = x^2\}$ ist ein volles Repräsentantensystem für \sim . Für welche Zahlen a, b, c ist die Abbildung $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ mit

$$f(x, y) = ax + by^2 + cx^2y^4 + dx^3y^3$$

\sim -invariant?

5. Für $x, y \in \mathbb{Z}$ sei $x \sim y$ genau dann, wenn $x - y$ durch 11 teilbar ist. Zeigen Sie, dass \sim eine Äquivalenzrelation auf \mathbb{Z} ist und dass $M = \{0\} \cup \{2^k \mid k = 0, \dots, 9\}$ ein volles Repräsentantensystem von \sim ist.

6 Körper

In diesem Abschnitt sollen die Regeln der Addition und der Multiplikation von Zahlen auf eine axiomatische Grundlage gestellt werden.

Aus bekannten Rechenregeln für die Addition und Multiplikation reeller Zahlen kann man neue Regeln durch logisches Schließen ableiten. Dabei bemerkt man, dass es nicht darauf ankommt, dass die Elemente, mit denen man rechnet, Zahlen sind oder eine andere konkrete Bedeutung haben.

Daher betrachtet man anstelle der Menge \mathbb{R} der reellen Zahlen zunächst eine beliebige Menge K und fordert, dass auf K zwei Rechenoperationen, die wir Addition und Multiplikation nennen, gegeben sind. Formal ausgedrückt bedeutet das: Es sind zwei Abbildungen

$$\begin{aligned} + : K \times K &\rightarrow K, & (x, y) &\mapsto x + y & \text{ und} \\ \cdot : K \times K &\rightarrow K, & (x, y) &\mapsto x \cdot y \end{aligned}$$

gegeben. Das Bild $x + y$ von (x, y) unter der Abbildung $+$ heißt die Summe von x und y , und das Bild $x \cdot y$, das wir wie gewohnt auch kurz als xy schreiben, heißt das Produkt von x und y .

Die von uns als grundlegend erkannten Rechenregeln in \mathbb{R} - es sollten möglichst wenige sein, aber doch so viele, dass die bekannten Rechenregeln in \mathbb{R} aus ihnen ableitbar sind - postulieren wir nun als Axiome für das Rechnen mit den Rechenoperatoren $+$ und \cdot in K . Diese Axiome sind dann der Ausgangspunkt für eine mathematische Theorie; in diesem Fall wird es die *Körpertheorie* sein. Die Axiome selbst sind die Grundsätze der Theorie und müssen nicht bewiesen werden. Aus den Axiomen werden die Lehrsätze der Theorie abgeleitet.

Axiomensysteme dürfen keinesfalls willkürlich aufgestellt werden. Erst nach guter Kenntnis eines Modells, wie in unserem Fall \mathbb{R} mit $+$ und \cdot , entwickelt sich ein abstraktes Axiomensystem.

Eine konkrete mathematische Struktur, die die Axiome erfüllt, heißt ein Modell für diese Axiome. Die einmal aus den Axiomen hergeleiteten Sätze gelten dann in *jedem* Modell für diese Axiome.

Wir werden jetzt die so genannten Körperaxiome aufstellen. Jedes Modell für diese Axiome nennt man dann einen Körper. Neben unseren Beispielen \mathbb{R} und \mathbb{Q} werden wir unendlich viele weitere Körper kennen lernen. Allen ist aber gemeinsam, dass man in ihnen im wesentlichen wie in \mathbb{R} und \mathbb{Q} rechnen kann. Auch was wir bisher über reelle lineare Gleichungssysteme gelernt haben, gilt genauso für lineare Gleichungssysteme in n Unbekannten, deren Koeffizienten und konstante Terme Elemente eines Körpers sind. Die Lösungen solcher Gleichungssysteme sind dann n -Tupel von Körperelementen.

Definition 6.1 Ein Körper ist eine Menge K zusammen mit einer **Addition** $+$: $K \times K \rightarrow K$, $(x, y) \mapsto x + y$ und einer **Multiplikation** \cdot : $K \times K \rightarrow K$, $(x, y) \mapsto xy$, so dass die folgenden Axiome (die Körperaxiome) erfüllt sind.

- (K1) Für alle $x, y, z \in K$ gilt: $(x + y) + z = x + (y + z)$.
(Die Addition ist **assoziativ**.)

- (K2) Für alle $x, y \in K$ gilt: $x + y = y + x$.
(Die Addition ist **kommutativ**.)
- (K3) Es gibt ein Element $0 \in K$ mit $x + 0 = x$ für alle $x \in K$.
(0 ist eindeutig bestimmt, denn ist auch $0' + x = x$ für alle $x \in K$, so gilt:
 $0 = 0' + 0 = 0 + 0' = 0'$. 0 heißt das **Nullelement** von K , kurz: **Null**.)
- (K4) Zu jedem $x \in K$ gibt es ein Element $-x \in K$ mit $x + (-x) = 0$.
($-x$ ist eindeutig bestimmt, denn gilt: $x + x' = 0$ für ein Element $x' \in K$, so folgt

$$x' = x' + 0 = x' + (x + (-x)) \stackrel{(K1)}{=} (x' + x) + (-x) \stackrel{(K2)}{=} (x + x') + (-x) = 0 + (-x) \stackrel{(K2)}{=} (-x) + 0 \stackrel{(K3)}{=} -x.$$
 $-x$ heißt das **negative Element** oder auch das **additive Inverse** von x .)
- (K5) Für alle $x, y, z \in K$ gilt: $(xy)z = x(yz)$.
(Die Multiplikation ist **assoziativ**.)
- (K6) Für alle $x, y \in K$ gilt: $xy = yx$.
(Die Multiplikation ist **kommutativ**.)
- (K7) Es gibt ein Element $1 \in K \setminus \{0\}$ mit $1x = x$ für alle $x \in K$.
(1 ist eindeutig bestimmt, denn ist $1' \in K$ mit $1'x = x$ für alle $x \in K$, so gilt:
 $1 = 1'1 = 11' = 1'$. 1 heißt das **Einselement** von K , kurz: **Eins**.)
- (K8) Zu jedem Element $x \in K \setminus \{0\}$ gibt es ein Element $x^{-1} \in K$ mit $x^{-1}x = 1$.
(x^{-1} ist eindeutig bestimmt, denn gilt: $x'x = 1$ für ein Element $x' \in K$, so folgt

$$x' = 1x' = (x^{-1}x)x' = x^{-1}(xx') = x^{-1}(x'x) = x^{-1}1 = 1x^{-1} = x^{-1}.$$
 x^{-1} heißt das **Inverse** von x oder auch das **Reziproke** von x und wird manchmal auch mit $\frac{1}{x}$ bezeichnet.)
- (K9) Für alle $x, y, z \in K$ gilt: $x(y + z) = xy + xz$.
(Es gilt das **Distributivgesetz**.)
Wegen (K6) gilt natürlich auch $(x + y)z = xz + yz$.

Bemerkung 6.2 In Definition 6.1 müsste es eigentlich präzise heißen: $(K, +, \cdot)$ ist ein Körper. Wenn es uns wichtig erscheint, zu betonen, dass zu einem Körper neben der Menge K auch die Rechenoperationen $+, \cdot$ gehören, wählen wir die Bezeichnung $(K, +, \cdot)$ anstelle der kurzen Bezeichnung K .

Notationen 6.3 Ist K ein Körper, und sind $x, y \in K$, so schreibt man kurz $x - y$ statt $x + (-y)$ und auch $\frac{x}{y}$ statt xy^{-1} oder $y^{-1}x$, falls $y \neq 0$. Also:

$$x - y := x + (-y) \quad \text{und} \quad \frac{x}{y} := xy^{-1}, \text{ falls } y \neq 0.$$

Lemma 6.4 Es sei K ein Körper. Dann gilt:

- a) $0 \cdot a = 0$ für alle $a \in K$.

b) Für alle $a, b \in K$ gilt: $ab = 0 \Leftrightarrow a = 0$ oder $b = 0$.
(Man sagt: K ist **nullteilerfrei**.)

c) Für alle $a, b \in K$ gilt:

$$(-a)b = -(ab), \quad (-a)(-b) = ab, \quad -a = (-1)a.$$

Beweis: zu a): $0a = (0 + 0)a = 0a + 0a$, also

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

zu b): Sei $ab = 0$ und $a \neq 0$. Dann ist

$$b = (a^{-1}a)b = a^{-1}(ab) = a^{-1}0 = 0a^{-1} = 0.$$

zu c):

i) $a + (-a) = 0 \Rightarrow ab + (-a)b = (a + (-a))b = 0b = 0 \Rightarrow -(ab) = (-a)b.$

ii) $(-a)(-b) = -(a(-b)) = -(-(ab)) = ab,$
denn: $-(-x) = x$, weil $x + (-x) = 0 = (-(-x)) + (-x).$

iii) $-a = -(a1) = a(-1).$ □

Lemma 6.5 In einem Körper K gelten die folgenden Regeln. Es seien $a, b, c, d \in K$, $b, d \neq 0$.

a)

$$\frac{a}{b} = \frac{c}{d} \iff ad = bc$$

b)

$$\frac{a}{b} + \frac{c}{d} = \frac{ad + bc}{bd}$$

c)

$$-\frac{a}{b} = \frac{-a}{b} = \frac{a}{-b}$$

d)

$$\frac{a}{b} \cdot \frac{c}{d} = \frac{ac}{bd}$$

e)

$$\left(\frac{a}{b}\right)^{-1} = \frac{b}{a}, \quad \text{falls auch } a \neq 0.$$

Der Beweis ist eine einfache Übung.

b) kann man z.B. so beweisen:

$$\begin{aligned} \frac{a}{b} + \frac{c}{d} &= ab^{-1} + cd^{-1} = ab^{-1}dd^{-1} + cbb^{-1}d^{-1} \\ &= ad(bd)^{-1} + bc(bd)^{-1} = (ad + bc)(bd)^{-1} = \frac{ad + bc}{bd}. \end{aligned}$$

Eine andere Möglichkeit wäre: $x = \frac{a}{b}$ ist die Lösung der Gleichung $bx = a$ und, $y = \frac{c}{d}$ ist die Lösung der Gleichung $dy = c$. Es folgt dann $bdx = ad$ und $bdy = bc$ und somit $bd(x + y) = ad + bc$, also $x + y = \frac{ad+bc}{bd}$. \square

Beispiel 6.6

a) $(\mathbb{Q}, +, \cdot)$ und $(\mathbb{R}, +, \cdot)$ sind Körper.

b) Sei $K_2 = \{a + b\sqrt{2} \mid a, b \in \mathbb{Q}\}$. K_2 ist eine Teilmenge von \mathbb{R} . Sind nun $x = a + b\sqrt{2}$ und $y = c + d\sqrt{2}$ Elemente in K_2 , so gilt, weil ja \mathbb{R} ein Körper ist: $x + y = a + c + (b + d)\sqrt{2}$. Da $a, b, c, d \in \mathbb{Q}$ sind, sind auch $a + c$ und $b + d$ Elemente in \mathbb{Q} , d.h. $x + y$ ist Element von K_2 . Die Einschränkung der Addition $+: \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$ auf $K_2 \times K_2$ ergibt also eine Addition $+: K_2 \times K_2 \rightarrow K_2$. Ebenso ergibt sich $xy \in K_2$, denn:
 $xy = (a + b\sqrt{2})(c + d\sqrt{2}) = ac + 2bd + (ad + bc)\sqrt{2}$ und $ac + 2bd, ad + bc \in \mathbb{Q}$.
 Die Multiplikation auf \mathbb{R} induziert also eine Multiplikation

$$\cdot : K_2 \times K_2 \rightarrow K_2.$$

Weiter gilt: $0, 1 \in K_2$, $-x = (-a) + (-b)\sqrt{2} \in K_2$, und schließlich ist auch

$$x^{-1} = \frac{1}{a + b\sqrt{2}} = \frac{a - b\sqrt{2}}{a^2 - 2b^2} = \frac{a}{a^2 - 2b^2} + \frac{-b}{a^2 - 2b^2}\sqrt{2} \in K_2,$$

falls $x = a + b\sqrt{2} \neq 0$ ist (Beweise dazu: Aus $a, b \in \mathbb{Q}$ und $a + b\sqrt{2} \neq 0$ folgt $a^2 - 2b^2 \neq 0$). Alle Axiome (1) - (9) sind für $(K_2, +, \cdot)$ erfüllt. Man nennt K_2 (wie auch \mathbb{Q}) einen Unterkörper von \mathbb{R} .

Allgemein definiert man

Definition 6.7 Es sei K ein Körper und $F \subset K$ eine Teilmenge. F heißt **Unterkörper** von K , wenn gilt:

- Für alle $x, y \in F$ gilt: $x - y \in F$.
- $1 \in F$.
- Für alle $x, y \in F$ ist $xy \in F$.
- Für alle $x \in F \setminus \{0\}$ ist $x^{-1} \in F$.

Offensichtlich ist $(F, +, \cdot)$ dann ein Körper (Beweis als Übung). Ist F ein Unterkörper von K , so nennt man K auch **Erweiterungskörper** von F .

Beispiel 6.8 In einem Körper kann man zwar im wesentlichen rechnen wie mit Zahlen, aber nicht alle Körper sind Erweiterungskörper von \mathbb{Q} . Nach den Körperaxiomen muss ein Körper stets mindestens zwei verschiedene Elemente enthalten, nämlich 0 und 1. Verblüffend ist nun, dass man auf einer Menge mit zwei Elementen eine Addition und eine Multiplikation definieren kann, so dass ein Körper entsteht: $(\mathbb{F}_2, +, \cdot)$ mit $\mathbb{F}_2 = \{G, U\}$ und den Verknüpfungstabellen ($G = \text{gerade}, U = \text{ungerade}$).

$$\begin{array}{c|cc} + & G & U \\ \hline G & G & U \\ U & U & G \end{array} \quad \begin{array}{c|cc} \cdot & G & U \\ \hline G & G & G \\ U & G & U \end{array}$$

ist ein Körper. G ist das Nullelement: $G = 0$ und, U ist das Einselement: $U = 1$. In diesem Körper gilt: $1 + 1 = 0$.

Definition 6.9 Es sei K ein Körper. Zur Unterscheidung von $0, 1 \in \mathbb{Z}$ sei 0_K das Nullelement und 1_K das Einselement in K . Die Abbildung

$$\mathbb{Z} \times K \rightarrow K, \quad (n, x) \mapsto nx$$

wird folgendermaßen definiert:

$$0x := 0_K \text{ für alle } x \in K,$$

$$nx := (n-1)x + x \text{ für alle } n \in \mathbb{Z}, n > 0, x \in K;$$

und schließlich

$$nx := -((-n)x) \text{ für alle } n \in \mathbb{Z}, n < 0, x \in K.$$

Offensichtlich gelten dann (Beweis als Übung) die Regeln

$$(n+m)x = nx + mx \text{ für } n, m \in \mathbb{Z}, x \in K$$

$$(nm)(xy) = (nx)(my) \text{ für } n, m \in \mathbb{Z}, x, y \in K.$$

Ist die Abbildung

$$\varphi: \mathbb{Z} \rightarrow K \text{ mit } \varphi(n) := n1_K$$

injektiv, so heißt K ein **Körper der Charakteristik Null**. φ lässt sich dann fortsetzen zu einer injektiven Abbildung $\psi: \mathbb{Q} \rightarrow K$. Man setzt

$$\psi\left(\frac{a}{b}\right) := \frac{a \cdot 1_K}{b \cdot 1_K}$$

für alle $a, b \in \mathbb{Z}, b \neq 0$.

ψ heißt die kanonische Abbildung von \mathbb{Q} in K .

Ist die Abbildung $\varphi: \mathbb{Z} \rightarrow K$ nicht injektiv, so gibt es ein $m \in \mathbb{Z}$ mit $m > 0$ und $m1_K = 0$, wie man sofort sieht. Es sei $p \in \mathbb{Z}, p > 0$, die kleinste positive ganze Zahl mit $p \cdot 1_K = 0_K$.

K heißt dann ein **Körper der Charakteristik p** . $\text{char}(K) = p$ heißt die **Charakteristik** von K . Man bemerkt sofort, dass p eine Primzahl ist. Ist nämlich $p = p_1 p_2$ mit $p_1, p_2 \in \mathbb{N}$, so ist $0_K = (p_1 p_2)1_K = (p_1 1_K)(p_2 1_K)$ also ohne Einschränkung $p_1 1_K = 0$. Da p minimal gewählt war, folgt $p = p_1, p_2 = 1$. p ist also unzerlegbar. Außerdem ist $p > 1$, weil sonst $1_K = 1 \cdot 1_K = 0_K$ wäre. p ist also eine Primzahl.

Lemma 6.10 Ist K ein Körper der Charakteristik $p > 0$, so ist

$$F = \{0_K, 1_K, 2 \cdot 1_K, \dots, (p-1) \cdot 1_K\}$$

ein Unterkörper von K .

Beweis: Man bemerkt zunächst, dass $n1_K = m1_K$, falls $n = qp + m$ für ein $q \in \mathbb{Z}$, (denn: $(qp)1_K = q \cdot (p \cdot 1_K) = q \cdot 0_K = 0_K$).

Zu jedem $n \in \mathbb{Z}$ gibt es also genau ein m mit $0 \leq m < p$, so dass $n1_K = m1_K$. Damit folgt leicht, dass a), b), c) in Definition 3.1.7 erfüllt sind.

Ist nun $x \in F$, $x \neq 0$, so ist die Abbildung $f : F \rightarrow F$ mit $f(y) := xy$ injektiv, denn aus $y_1 \neq y_2$ folgt $y_1 - y_2 \neq 0$, also auch $x(y_1 - y_2) \neq 0$ und somit $xy_1 \neq xy_2$. Da F eine endliche Menge ist, ist f somit bijektiv, und es gibt insbesondere ein Element $y \in F$ mit $xy = 1$, d.h. $x^{-1} \in F$. Damit ist gezeigt, dass F ein Unterkörper von K ist. \square

Jetzt wollen wir zeigen, dass es zu jeder Primzahl p einen Körper der Charakteristik p gibt.

Beispiel 6.11 Es sei eine Primzahl p fest gegeben. Mit \mathbb{F}_p bezeichnen wir eine Menge aus p Elementen. Den Elementen von \mathbb{F}_p geben wir die Namen

$$[0], [1], \dots, [p-1].$$

Sodann definieren wir die Abbildung

$$\begin{aligned} \mathbb{Z} &\longrightarrow \mathbb{F}_p \\ n &\longmapsto n \bmod p := [r], \end{aligned}$$

wobei r der Rest beim Teilen von n durch p ist, d.h. r ist diejenige natürliche Zahl mit $0 \leq r < p$, für die gilt: $\exists q \in \mathbb{Z}$, so dass $n = qp + r$.

Im Fall $p = 13$ gilt zum Beispiel:

$$\begin{aligned} 22 \bmod 13 &= [9] \\ 26 \bmod 13 &= [0] \\ 30 \bmod 13 &= [4] \quad \text{usw..} \end{aligned}$$

Auf der Menge \mathbb{F}_p erklären wir eine Addition, die wir zur Unterscheidung von $+$ auf \mathbb{Z} vorläufig mit \oplus bezeichnen:

Für $0 \leq r_1 < p$, $0 \leq r_2 < p$ sei $[r_1] \oplus [r_2] := [r]$, wobei r der Rest beim Teilen von $r_1 + r_2$ durch p ist. In Formeln:

$$[r_1] \oplus [r_2] = (r_1 + r_2) \bmod p.$$

Beispiel: In \mathbb{F}_{13} gilt: $[8] \oplus [7] = (8 + 7) \bmod 13 = [2]$.

Wir zeigen: Für alle $n, m \in \mathbb{Z}$ gilt:

$$n \bmod p \oplus m \bmod p = (n + m) \bmod p. \quad (*)$$

Beweis: Es seien $q, t, r, s \in \mathbb{Z}$, $0 \leq r < p$, $p \leq s < 2p$, so dass $n = qp + r$, $m = tp + s$. Dann gilt:

$$n + m = (q + t)p + (r + s),$$

und somit

$$n \bmod p \oplus m \bmod p = [r] \oplus [s] = (r + s) \bmod p = (n + m) \bmod p.$$

Jetzt erklären wir eine Multiplikation \odot auf \mathbb{F}_p :

Für $0 \leq r_1 < p$, $0 \leq r_2 < p$ sei

$$[r_1] \odot [r_2] := [r],$$

wobei r der Rest beim Teilen von $r_1 r_2$ durch p ist. In Formeln:

$$[r_1] \odot [r_2] = (r_1 r_2) \bmod p.$$

Analog zu (*) gilt für alle $n, m \in \mathbb{Z}$:

$$n \bmod p \odot m \bmod p = (nm) \bmod p. \quad (**)$$

Beweis zu (**): Es sei wieder $n = qp + r$, $m = tp + s$, $n \bmod p = [r]$, $m \bmod p = [s]$. Dann gilt: $nm = (qt + p + rt + sq)p + rs$, also

$$nm \bmod p = rs \bmod p,$$

und somit

$$n \bmod p \odot m \bmod p = [r] \odot [s] = (rs) \bmod p = (nm) \bmod p.$$

Jetzt können wir zeigen:

\mathbb{F}_p ist mit der Addition \oplus und Multiplikation \odot ein Körper.

Beweis: Da $(\mathbb{Z}, +, \cdot)$ die Axiome (K1) - (K7) und (K9) erfüllt, kann man mit Hilfe von (*) und (**) folgern, dass $(\mathbb{F}_p, \oplus, \odot)$ ebenfalls die Axiome (K1) - (K7) und (K9) erfüllt.

$[0]$ ist das Nullelement und $[1]$ ist das Einselement in \mathbb{F}_p . Weiter gilt: $[p - r]$ ist das negative Element von $[r]$ für $0 < r < p$.

Die einzige Aufgabe, die noch bleibt, ist die Verifikation von Axiom (K8).

Wir zeigen: Zu jeder natürlichen Zahl r mit $0 < r < p$ gibt es eine natürliche Zahl s mit $0 < s < p$, so dass

$$[r] \odot [s] = [1]$$

gilt.

Zunächst zeigen wir:

Behauptung: Ist $0 \leq s < p$ und $[r] \odot [s] = [0]$, so ist $s = 0$.

Beweis: Ist $[r] \odot [s] = [0]$, so ist p Teiler von rs . Da p kein Teiler von r ist, muß p ein Teiler von s sein und da $0 \leq s < p$ ist, folgt $s = 0$.

Weiter zeigen wir:

Behauptung: Die Abbildung $f : \mathbb{F}_p \rightarrow \mathbb{F}_p$ mit $f(x) = [r] \odot x$ ist bijektiv.

Beweis: Ist $f(x_1) = f(x_2)$, so ist $[r] \odot (x_1 \ominus x_2) = f(x_1 \ominus x_2) = f(x_1) \ominus f(x_2) = [0]$, also nach a). $x_1 \ominus x_2 = [0]$, d.h. $x_1 = x_2$. Also ist f injektiv. Da \mathbb{F}_p endlich ist, ist f bijektiv.

Insbesondere ist $[1]$ im Bild von f , d.h. es gibt ein $[s] \in \mathbb{F}_p$ mit

$$[r] \odot [s] = [1].$$

Damit ist bewiesen, dass \mathbb{F}_p ein Körper ist.

Beispiel 6.12 In \mathbb{F}_{13} gilt:

$$\begin{aligned} [3]^{-1} &= \frac{1}{[3]} = \frac{1 \bmod 13}{3 \bmod 13} = \frac{-12 \bmod 13}{3 \bmod 13} = -4 \bmod 13 = [9], [3] \odot [9] \\ &= 27 \bmod 13 = [1], \\ [5]^{-1} &= \frac{1}{[5]} = \frac{1 \bmod 13}{5 \bmod 13} = \frac{5 \bmod 13}{25 \bmod 13} = \frac{5 \bmod 13}{-1 \bmod 13} = -5 \bmod 13 = [8], [5] \odot [8] \\ &= 40 \bmod 13 = [1] \end{aligned}$$

usw..

Ein Algorithmus zur Bestimmung des Inversen wird durch den so genannten **erweiterten euklidischen Algorithmus** gegeben. Um $[r] \odot [s] = [1]$ zu lösen, muss man nämlich zu gegebenen r, p die ganzzahligen Lösungen s, t der Gleichung

$$rs + pt = 1$$

bestimmen, und das leistet der genannte Algorithmus (siehe [23]).

Definition 6.13 Es seien K und L Körper. Eine Abbildung $f : K \rightarrow L$ heißt Körperisomorphismus, wenn f die folgenden Eigenschaften hat:

- a) f ist bijektiv.
- b) $f(x + y) = f(x) + f(y)$ für alle $x, y \in K$.
- c) $f(1) = 1$.
- d) $f(xy) = f(x)f(y)$ für alle $x, y \in K$.

Es gilt dann auch $f(0) = 0$, denn $f(0) = f(0 + 0) = f(0) + f(0)$. Weiter gilt: $f(-x) = -f(x)$, denn

$$0 = f(0) = f(x + (-x)) = f(x) + f(-x).$$

Schließlich gilt auch: $f(x) \neq 0$, falls $x \neq 0$ und $f(x)^{-1} = f(x^{-1})$. Ist $K = L$, so heißt f **Körperautomorphismus** von K . Zwei Körper K und L heißen **isomorph** (in Zeichen: $K \cong L$), wenn es einen Körperisomorphismus $f : K \rightarrow L$ gibt.

Beispiel 6.14 Ist K ein Körper der Charakteristik $p > 0$ und 1_K das Einselement von K , F der Unterkörper $\{0 \cdot 1_K, 1 \cdot 1_K, 2 \cdot 1_K, \dots, (p-1) \cdot 1_K\}$, so ist $F \cong \mathbb{F}_p$, und zwar ist

$$f : \mathbb{F}_p \rightarrow F \text{ mit } f([r]) := r1_K$$

ein Isomorphismus (Übung).

Ist K ein Körper der Charakteristik Null, so enthält K einen zu \mathbb{Q} isomorphen Unterkörper F , nämlich $\psi(\mathbb{Q})$, wobei $\psi : \mathbb{Q} \rightarrow K$ die kanonische Abbildung ist.

Jetzt kommen wir zu unserem wichtigsten Beispiel, dem Körper der komplexen Zahlen.

Satz 6.15 Es gibt einen Körper $(\mathbb{C}, +, \cdot)$ mit folgenden Eigenschaften:

- a) $\mathbb{R} \subset \mathbb{C}$ ist ein Unterkörper von \mathbb{C} .
- b) Es gibt ein Element $i \in \mathbb{C}$ mit $i^2 = -1$.
- c) Zu jedem Element $z \in \mathbb{C}$ gibt es eindeutig bestimmte reelle Zahlen x und y , so dass

$$z = x + yi.$$

Durch a) - c) ist \mathbb{C} bis auf Isomorphie eindeutig bestimmt.

Beweis:

1. Eindeutigkeit: Aus a) - c) folgt:

$$f : \mathbb{C} \rightarrow \mathbb{R}^2, \quad z = x + yi \mapsto (x, y)$$

ist eine bijektive Abbildung. Da \mathbb{C} ein Körper ist, gilt für $z = x + yi$, $w = u + vi$ nach den Körperaxiomen:

$$z + w = (x + u) + (y + v)i,$$

also

$$f(z + w) = (x + u, y + v),$$

und wegen b)

$$zw = (xu - yv) + (xv + yu)i,$$

also

$$f(zw) = (xu - yv, xv + yu).$$

f ist also ein Körperisomorphismus von \mathbb{C} auf \mathbb{R}^2 , wenn man \mathbb{R}^2 mit der folgenden Addition und Multiplikation versieht:

$$\left. \begin{aligned} (x, y) + (u, v) &:= (x + u, y + v) \\ (x, y) \cdot (u, v) &:= (xu - yv, xv + yu) \end{aligned} \right\} (*)$$

Wenn es einen Körper \mathbb{C} mit den Eigenschaften a) - c) gibt, so ist er also zu \mathbb{R}^2 mit den Verknüpfungen $(*)$ isomorph. Damit ist die Eindeutigkeit bewiesen.

2. Existenz: Wir müssen zeigen, dass \mathbb{R}^2 mit den Verknüpfungen $(*)$ die Körperaxiome (K1) - (K9) erfüllt. (K1) - (K4) sind trivialerweise erfüllt. $0 = (0, 0)$ ist das Nullelement in \mathbb{R}^2 , und $-(x, y) = (-x, -y)$ ist das Negative von (x, y) .

Die Axiome der Multiplikation sind auch einfach nachzuweisen, jedoch etwas länger. Wir zeigen z.B. (K5) und (K8) und lassen den Rest als Übung:

zu (K5):

$$\begin{aligned} ((x, y) \cdot (x', y')) \cdot (x'', y'') &= (xx' - yy', xy' + yx') \cdot (x'', y'') \\ &= ((xx' - yy')x'' - (xy' + yx')y'', (xx' - yy')y'' + (xy' + yx')x'') \\ &= (xx'x'' - yy'x'' - xy'y'' - yx'y'', xx'y'' - yy'y'' + xy'x'' + yx'x''). \end{aligned}$$

Jetzt berechnet man $(x, y) \cdot ((x', y') \cdot (x'', y''))$. Man kommt auf dasselbe Ergebnis.

zu (K8): Zunächst ist $1 := (1, 0)$ das Einselement, wie man durch Einsetzen in die Definition sieht. Es sei nun $(x, y) \in \mathbb{R}^2$, $(x, y) \neq 0$. Man erhält

$$(x, y)(x, -y) = (x^2 + y^2, 0),$$

und somit

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

Damit haben wir die Behauptung bewiesen. Komplexe Zahlen $z \in \mathbb{C}$ schreiben wir in der Form

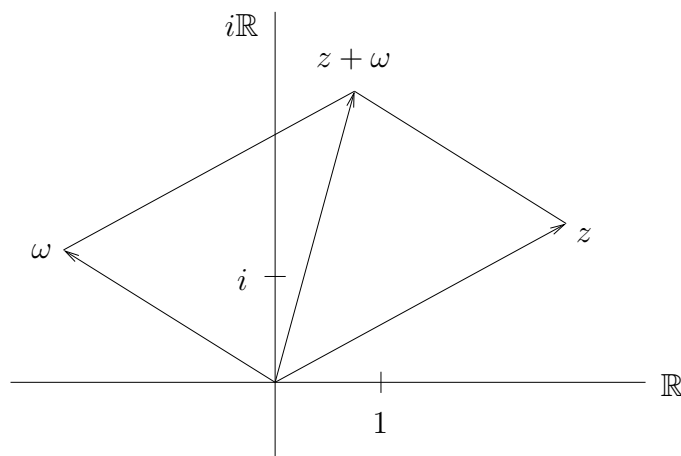
$$z = x + iy \text{ mit } x, y \in \mathbb{R}.$$

$x = \operatorname{Re}(z)$ heißt der **Realteil** von z , $y = \operatorname{Im}(z)$ heißt der **Imaginärteil** von z .

Durch die bijektive Abbildung

$$\begin{aligned} \mathbb{C} &\longrightarrow \mathbb{R}^2 \\ z &\longmapsto (\operatorname{Re}(z), \operatorname{Im}(z)) \end{aligned}$$

können wir uns \mathbb{C} als Ebene (die **komplexe Ebene**) vorstellen:



1 ist der Punkt $(1, 0) \in \mathbb{R}^2$ und, i ist der Punkt $(0, 1)$. Die x -Achse heißt dann die **reelle Achse** \mathbb{R} und die y -Achse heißt die **imaginäre Achse** $i\mathbb{R} = \{iy | y \in \mathbb{R}\}$.

Ist $z = x + iy \in \mathbb{C}$, so heißt

$$\bar{z} = x - iy$$

die **konjugiert komplexe Zahl** zu z .

Man zeigt leicht:

$$z \mapsto \bar{z} \text{ ist ein Körperautomorphismus } \mathbb{C} \rightarrow \mathbb{C},$$

und es gilt $z = \bar{z}$ genau dann, wenn z reell ist. Weiter ist

$$z\bar{z} = (x + iy)(x - iy) = x^2 - (iy)^2 = x^2 + y^2 \geq 0.$$

Daher nennt man

$$|z| := +\sqrt{z\bar{z}} \text{ den } \mathbf{Betrag} \text{ von } z.$$

Es sei z eine von Null verschiedene komplexe Zahl. Dann ist $r = |z|$ eine positive reelle Zahl und $u := \frac{z}{r}$ ist eine komplexe Zahl vom Betrag 1, denn

$$|u| = \left| \frac{z}{r} \right| = \frac{1}{r}|z| = 1.$$

Die komplexen Zahlen vom Betrag 1 werden auch **unimodulare** Zahlen genannt. Sie beschreiben den Einheitskreis

$$S^1 = \{x + iy \mid x^2 + y^2 = 1\} \subset \mathbb{C}$$

in der komplexen Ebene. Wie man leicht sieht, ist das Produkt unimodularer Zahlen wieder unimodular. Wir setzen nun die trigonometrischen Funktionen Cosinus und Sinus als bekannt voraus, zu mindestens die folgenden Aussagen:

- (i) \cos und \sin sind periodisch mit der Periode 2π .
- (ii) $\cos^2 \varphi + \sin^2 \varphi = 1$ für alle $\varphi \in \mathbb{R}$
- (iii) Zu jeder unimodularen Zahl $u = x + iy$ gibt es genau eine reelle Zahl $\varphi \in \mathbb{R}$ mit $0 \leq \varphi < 2\pi$, so dass

$$x = \cos \varphi \text{ und } y = \sin \varphi,$$

also

$$u = \cos \varphi + i \sin \varphi$$

gilt.

- (iv) Es gelten die Additionstheoreme für Sinus und Cosinus:

$$\begin{aligned} \cos(\alpha + \beta) &= \cos \alpha \cos \beta - \sin \alpha \sin \beta, \\ \sin(\alpha + \beta) &= \cos \alpha \sin \beta + \sin \alpha \cos \beta. \end{aligned}$$

Wir definieren nun für $\varphi \in \mathbb{R}$

$$e^{i\varphi} := \cos \varphi + i \sin \varphi. \quad (27)$$

Nach (iv) gilt dann

$$\begin{aligned} e^{i(\alpha+\beta)} &= \cos(\alpha + \beta) + i \sin(\alpha + \beta) \\ &= (\cos \alpha + i \sin \alpha)(\cos \beta + i \sin \beta) \\ &= e^{i\alpha} e^{i\beta} \end{aligned}$$

Damit haben wir nach (iii) eine surjektive Abbildung

$$\mathbb{R} \longrightarrow S^1, \quad \varphi \longmapsto e^{i\varphi}$$

mit den Eigenschaften

- a) $e^{i\varphi} = e^{i\psi} \iff \varphi - \psi \in 2\pi\mathbb{Z}$
- b) $e^{i(\alpha+\beta)} = e^{i\alpha} e^{i\beta}$

Jede von Null verschiedene komplexe Zahl z hat eine Darstellung der Form

$$z = r e^{i\varphi}$$

wobei $r = |z| > 0$ und $\varphi \in \mathbb{R}$.

φ ist bis auf ganzzahlige Vielfache von 2π durch z bestimmt und heißt ein **Argument** von z .

Man nennt $z = r e^{i\varphi}$ die **Polarkoordinatendarstellung** von z . Diese Darstellung erlaubt eine geometrische Interpretation der Multiplikation komplexer Zahlen.

Ist $z = r e^{i\varphi}$ und $w = s e^{i\psi}$, so ist $zw = r s e^{i(\varphi+\psi)}$.

Zwei komplexe Zahlen werden multipliziert, indem man ihre Beträge multipliziert und ihre Argumente addiert.

Die Bedeutung der komplexen Zahlen liegt in der Tatsache, dass jede algebraische Gleichung

$$x^n + a_{n-1}x^{n-1} + \dots + a_1x + a_0 = 0$$

mit $a_0, \dots, a_{n-1} \in \mathbb{C}$ ($n \geq 1$) im Körper der komplexen Zahlen lösbar ist. Es gibt n komplexe Zahlen (die auch mehrfach auftreten können) $z_1, \dots, z_n \in \mathbb{C}$, so dass

$$x^n - a_{n-1}x^{n-1} + \dots + a_0 = (x - z_1) \dots (x - z_n)$$

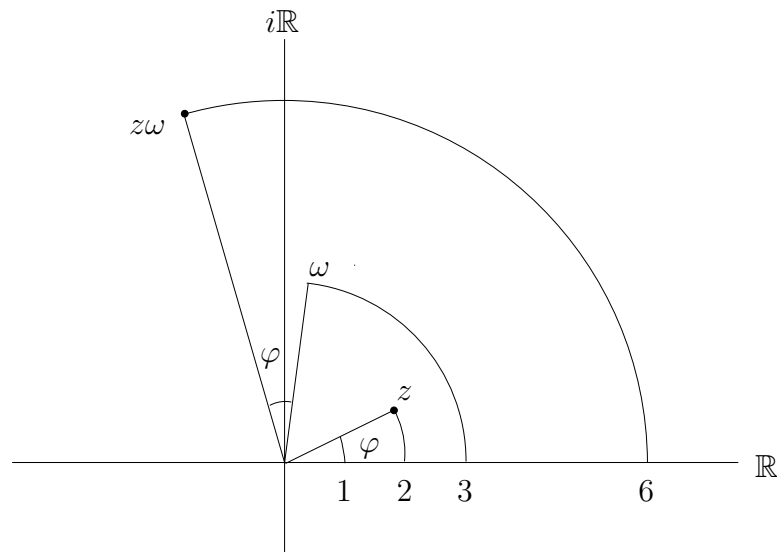
gilt. Dies ist der berühmte Fundamentalsatz der Algebra, den wir hier nicht beweisen werden.

Insbesondere kann man in \mathbb{C} n -te Wurzeln ziehen: Es sei $a = r e^{i\varphi} \in \mathbb{C}$, $a \neq 0$. Dann gibt es n komplexe Lösungen der Gleichung $x^n - a = 0$.

Eine Lösung ist

$$\omega_0 = \sqrt[n]{r} e^{i\frac{\varphi}{n}},$$

denn es gilt ja $\omega_0^n = (\sqrt[n]{r})^n e^{i\frac{\varphi}{n} \cdot n} = r e^{i\varphi} = a$.



Multiplikation komplexer Zahlen

Betrachtet man nun

$$\zeta = e^{\frac{2\pi i}{n}},$$

so gilt für $k = 0, \dots, n-1$:

$$(\zeta^k)^n = \zeta^{kn} = (\zeta^n)^k = (e^{2\pi i})^k = 1,$$

denn $e^{2\pi i} = \cos(2\pi) + i \sin(2\pi) = 1$.

Setzt man jetzt

$$\omega_k := \omega_0 \zeta^k = \sqrt[n]{r} e^{i \frac{\varphi + 2\pi k}{n}},$$

so hat man n verschiedene n -te Wurzeln von a : $\omega_0, \omega_1, \dots, \omega_{n-1}$, denn

$$\omega_k^n = \omega_0^n \zeta^{kn} = a \cdot 1 = a,$$

und $\omega_k = \omega_m \iff \zeta^k = \zeta^m \iff \zeta^{k-m} = 1 \iff k-m$ ist Vielfaches von n .

Übungen

1. Konstruieren Sie einen Körper K mit vier Elementen $0, 1, \alpha, \beta$, so dass $\{0, 1\}$ ein Unterkörper ist.
2. Finden Sie alle komplexen Zahlen $z \in \mathbb{C}$ mit $|z| = 1$, für die $z^2 + (1+i)z$ rein imaginär ist (d.h. der Realteil ist Null). Zeichnen Sie die Lösungsmenge in die komplexe Ebene.
3. Es sei $d \in \mathbb{Z}$ eine beliebige ganze Zahl und $\sqrt{d} \in \mathbb{C}$ eine Wurzel von d . Es gelte $\sqrt{d} \notin \mathbb{Q}$. Beweisen Sie:

$$\mathbb{Q}(\sqrt{d}) := \{a + b\sqrt{d} \mid a, b \in \mathbb{Q}\}$$

ist ein Unterkörper von \mathbb{C} . Wie viele Isomorphismen $f : \mathbb{Q}(\sqrt{d}) \rightarrow \mathbb{Q}(\sqrt{d})$ gibt es? Ist $\mathbb{Q}(\sqrt{2})$ zu $\mathbb{Q}(\sqrt{5})$ isomorph?

4. Es sei $z = \frac{2+i}{2-i}$. Berechnen Sie $|z|$. Ist die Menge $\{z^n | n \in \mathbb{Z}\}$ endlich?
5. Es sei $n \in \mathbb{N}$, $n \geq 2$. Eine n -te Einheitswurzel $\zeta \in \mathbb{C}$ heißt **primitiv**, wenn $\zeta^k \neq 1$ für $k = 1, \dots, n-1$. Berechnen Sie die Summe aller primitiven n -ten Einheitswurzeln für $n = 12, 13, 15$.
6. Aus der Gleichung

$$(\cos \varphi + i \sin \varphi)^n = \cos(n\varphi) + i \sin(n\varphi)$$

leite man die Formeln

$$\cos(n\varphi) = \sum_{0 \leq k \leq \frac{n}{2}} (-1)^k \binom{n}{2k} \cos^{n-2k} \varphi \cdot \sin^{2k} \varphi$$

und

$$\sin(n\varphi) = \sum_{0 \leq k \leq \frac{n-1}{2}} (-1)^k \binom{n}{2k+1} \cos^{n-1-2k} \varphi \cdot \sin^{2k+1} \varphi$$

ab. (\sum ist das Summenzeichen. Es wird über alle $k \in \mathbb{N}$ innerhalb der angegebenen Schranken summiert.)

7. Es sei p eine Primzahl. Beweisen Sie:
- (a) Für alle $a, b \in \mathbb{F}_p$ gilt:
- $$(a+b)^p = a^p + b^p.$$
- (b) Für alle $n \in \mathbb{Z}$ gilt: $n^p \bmod p = n \bmod p$, d.h. $n^p - n$ ist durch p teilbar.

7 Vektorräume

Es sei K ein Körper.

Definition 7.1 Ein K -Vektorraum ist eine Menge V , zusammen mit einer **Addition**

$$+ : V \times V \rightarrow V, \quad (x, y) \mapsto x + y$$

und einer Abbildung

$$\cdot : K \times V \rightarrow V, \quad (a, x) \mapsto ax,$$

die **skalare Multiplikation** genannt wird, so dass die folgenden Axiome (die Vektorraumaxiome) erfüllt sind:

(V1) $(x + y) + z = x + (y + z)$ für alle $x, y, z \in V$.

(V2) $x + y = y + x$ für alle $x, y \in V$.

(V3) Es gibt ein Element $0 \in V$ (**Null** oder **Nullvektor** genannt), so dass $x + 0 = x$ für alle $x \in V$.

(V4) Zu jedem Element $x \in V$ gibt es ein Element $-x \in V$ mit $x + (-x) = 0$.

(V5) $(ab)x = a(bx)$ für alle $a, b \in K, x \in V$.

(V6) $1x = x$ für alle $x \in V$.

(V7) $a(x + y) = ax + ay$ für alle $a \in K, x, y \in V$.

(V8) $(a + b)x = ax + bx$ für alle $a, b \in K, x \in V$.

Die Elemente von V heißen **Vektoren**.

Bemerkung 7.2 Die Axiome (V1) - (V4) für die Addition in V stimmen mit den Axiomen (K1) - (K4) für Körper überein. Auch in Vektorräumen gilt daher: Nullvektor und der negative Vektor $-x$ von x sind eindeutig durch (V3) bzw. (V4) bestimmt.

Beispiel 7.3 Es sei $n \in \mathbb{N}$. Dann ist

$$K^n = \{x \mid \exists x_1, \dots, x_n \in K, \text{ so dass } x = (x_1, \dots, x_n)\}$$

mit den folgenden Verknüpfungen ein K -Vektorraum.

Für $x = (x_1, \dots, x_n), y = (y_1, \dots, y_n) \in K^n$ und $a \in K$ sei

$$\begin{aligned} x + y &:= (x_1 + y_1, \dots, x_n + y_n) \\ ax &:= (ax_1, \dots, ax_n). \end{aligned}$$

Insbesondere ist \mathbb{R}^n ein \mathbb{R} -Vektorraum (**reeller Vektorraum**) und \mathbb{C}^n ist ein \mathbb{C} -Vektorraum (**komplexer Vektorraum**).

Beispiel 7.4 Es sei X eine beliebige Menge. $\text{Abb}(X, K)$ ist ein K -Vektorraum mit den folgenden Verknüpfungen: Für $f, g \in \text{Abb}(X, K)$, $a \in K$ sei

$$(f + g)(x) := f(x) + g(x) \text{ für alle } x \in X$$

und

$$(af)(x) := af(x) \text{ für alle } x \in X.$$

Lemma 7.5 Es sei V ein K -Vektorraum. Dann gilt:

- a) $0x = 0$ für alle $x \in V$.
- b) $a0 = 0$ für alle $a \in K$, $0 = \text{Nullvektor}$.
- c) $(-a)x = -(ax) = a(-x)$ für alle $a \in K$, $x \in V$.
- d) $-x = (-1)x$ für alle $x \in V$.
- e) Für alle $a \in K$, $x \in V$ gilt: Ist $ax = 0$, so ist $a = 0$ oder $x = 0$.
- f) Setzt man $x - y := x + (-y)$ für $x, y \in V$, so gilt:

$$a(x - y) = ax - ay \text{ für alle } a \in K, x, y \in V.$$

Beweis:

zu a): $x + 0x = 1x + 0x = (1 + 0)x = 1x = x$, also $0x = 0$.

zu b): $ax + a0 = a(x + 0) = ax$, also $a0 = 0$.

zu c): $ax + (-a)x = (a + (-a))x = 0x = 0$, also $(-a)x = -(ax)$ usw..

zu e): Es sei $ax = 0$ und $a \neq 0$. Dann ist $0 = a^{-1}0 = a^{-1}(ax) = (a^{-1}a)x = 1x = x$.

zu f): $a(x - y) = a(x + (-y)) = ax + a(-y) = ax + (-(ay)) = ax - ay$. \square

Definition 7.6 Es sei V ein K -Vektorraum. Eine Teilmenge $W \subset V$ heißt **Untervektorraum** von V , wenn folgende Bedingungen erfüllt sind:

- a) $W \neq \emptyset$.
- b) Für alle x, y gilt: Sind $x, y \in W$, so ist auch $x + y \in W$ (kurz: $x, y \in W \Rightarrow x + y \in W$).
- c) Für alle $a \in K$ gilt: Ist $x \in W$, so ist $ax \in W$ (kurz: $a \in K, x \in W \Rightarrow ax \in W$).

Bemerkung 7.7

- a) Offensichtlich ist eine nicht leere Teilmenge $W \subset V$ genau dann ein Untervektorraum von V , wenn gilt:

$$a, b \in K, x, y \in W \Rightarrow ax + by \in W.$$

- b) Ist $W \subset V$ ein Untervektorraum von V , so kann man die Addition $+: V \times V \rightarrow V$ und die skalare Multiplikation $\cdot: K \times V \rightarrow V$ auf W einschränken. Man erhält

$$+: W \times W \rightarrow W, \quad \cdot: K \times W \rightarrow W.$$

Mit diesen Operationen ist dann W ein K -Vektorraum. Dazu muss man nur zeigen, dass $0 \in W$ und, dass $-x \in W$, falls $x \in W$.

Beweis: Da $W \neq \emptyset$ ist, kann man ein $x \in W$ finden. Dann ist $-x = (-1)x \in W$ und auch $0 = x + (-x) \in W$. \square

Satz 7.8 Es sei V ein K -Vektorraum, $I \neq \emptyset$ sei eine Indexmenge, und W_i sei ein Untervektorraum von V für $i \in I$. Dann ist auch

$$W = \bigcap_{i \in I} W_i$$

ein Untervektorraum von V .

Beweis: Da $I \neq \emptyset$ ist, ist $W \neq \emptyset$. Seien nun $a, b \in K$ und $x, y \in W$. Dann gilt: $x, y \in W_i$ für alle $i \in I$. Da W_i Untervektorraum von V ist, folgt $ax + by \in W_i$ für alle $i \in I$. Also gilt: $ax + by \in W$. \square

Beachte: Die Vereinigung von Untervektorräumen ist nur in Ausnahmefällen ein Untervektorraum.

Definition 7.9 Es sei V ein K -Vektorraum, $n \in \mathbb{N}$. Weiter seien $x_1, \dots, x_n \in V$, $a_1, \dots, a_n \in K$. Dann heißt

$$x = a_1x_1 + a_2x_2 + \dots + a_nx_n$$

eine **Linearkombination** von x_1, \dots, x_n .

Wir verwenden auch das Summenzeichen und schreiben

$$x = \sum_{\nu=1}^n a_\nu x_\nu.$$

Mit $\langle x_1, \dots, x_n \rangle$ bezeichnen wir die Menge aller Linearkombinationen von x_1, \dots, x_n . Also,

$$\langle x_1, \dots, x_n \rangle = \{x \in V \mid \exists a_1, \dots, a_n \in K, \text{ so dass } x = \sum_{\nu=1}^n a_\nu x_\nu\}.$$

Ist $n = 1$, so schreiben wir auch $Kx_1 = \langle x_1 \rangle = \{ax_1 \mid a \in K\}$.

Satz 7.10 Es sei V ein K -Vektorraum, und es seien $x_1, \dots, x_n \in V$. Dann ist $\langle x_1, \dots, x_n \rangle$ ein Untervektorraum von V , und zwar der kleinste Untervektorraum von V , der x_1, \dots, x_n erhält.

Beweis:

- a) $\langle x_1, \dots, x_n \rangle \neq \emptyset$, weil $0 \in \langle x_1, \dots, x_n \rangle$ (Im Fall $n = 0$ wird $\langle x_1, \dots, x_n \rangle$ als Nullvektorraum $\{0\}$ definiert.).

- b) Sind $a, b \in K$ und $x = \sum_{\nu=1}^n a_{\nu}x_{\nu}$, $y = \sum_{\nu=1}^n b_{\nu}x_{\nu} \in \langle x_1, \dots, x_n \rangle$, so gilt: $ax + by = \sum_{\nu=1}^n (aa_{\nu} + bb_{\nu})x_{\nu} \in \langle x_1, \dots, x_n \rangle$. Damit ist $\langle x_1, \dots, x_n \rangle$ ein Untervektorraum von V .
- c) Ist $W \subset V$ ein Untervektorraum mit $x_1, \dots, x_n \in W$, so ist auch jede Linearkombination von x_1, \dots, x_n Element von W (Beweis durch Induktion nach n), also gilt: $\langle x_1, \dots, x_n \rangle \subset W$. \square

Definition 7.11 Es sei V ein K -Vektorraum und $M \subset V$ eine beliebige Teilmenge. Dann heißt der Untervektorraum (siehe Satz 7.8)

$$\text{Span}(M) := \bigcap_{W \text{ Untervektorraum von } V \text{ mit } M \subset W} W,$$

der von M **aufgespannte Untervektorraum** von V . Wir schreiben auch $\text{Span}(M) = \text{Span}_K(M)$.

Es gilt:

$$\begin{aligned} \text{Span}(\emptyset) &= \{0\}, \text{ und nach Satz 7.10 gilt:} \\ \text{Span}(\{x_1, \dots, x_n\}) &= \langle x_1, \dots, x_n \rangle \end{aligned}$$

für $x_1, \dots, x_n \in V$. Wir schreiben auch $\text{Span}(\{x_1, \dots, x_n\}) = \text{Span}(x_1, \dots, x_n) = \text{Span}_K(x_1, \dots, x_n)$.

$M \subset V$ heißt **Erzeugendensystem** von V , wenn

$$\text{Span}(M) = V$$

gilt. Ein n -Tupel (x_1, \dots, x_n) von Vektoren $x_1, \dots, x_n \in V$ heißt Erzeugendensystem von V , wenn $\{x_1, \dots, x_n\}$ ein Erzeugendensystem von V ist, d.h. wenn

$$\langle x_1, \dots, x_n \rangle = V$$

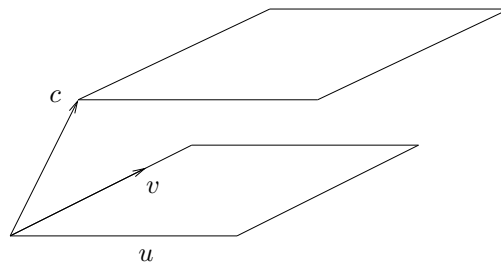
gilt. In diesem Fall heißt V **endlich erzeugt**.

Satz 7.12 Es sei V ein K -Vektorraum und $M \subset V$ eine nicht leere Teilmenge. Dann gilt für $x \in V$:

$$x \in \text{Span}(M) \Leftrightarrow \exists x_1, \dots, x_n \in M, a_1, \dots, a_n \in K, \text{ so dass} \\ x = a_1x_1 + \dots + a_nx_n \text{ gilt.}$$

Beweis:

$$W_0 := \{x \in V \mid \exists x_1, \dots, x_n \in M, a_1, \dots, a_n \in K : x = \sum_{\nu=1}^n a_{\nu}x_{\nu}\}$$

Parallele Ebenen $\langle u, v \rangle$ und $\langle u, v \rangle + c$

ist ein Untervektorraum von V mit $M \subset W_0$, denn $W_0 \neq \emptyset$ und für $a, b \in K$, $x = \sum_{\nu=1}^n a_{\nu}x_{\nu}$, $y = \sum_{\mu=1}^m b_{\mu}y_{\mu}$; $x_1, \dots, x_n, y_1, \dots, y_m \in M$, ist

$$ax + by = \sum_{\nu=1}^n (aa_{\nu})x_{\nu} + \sum_{\mu=1}^m (bb_{\mu})y_{\mu} \in W_0.$$

Offensichtlich gilt auch: $W_0 \subset W$ für jeden Untervektorraum W von V mit $M \subset W$.
□

Beispiel 7.13 a) Es seien $e_1 = (1, 0, 0)$, $e_2 = (0, 1, 0)$, $e_3 = (0, 0, 1) \in \mathbb{R}^3$. Dann gilt für alle: $x = (x_1, x_2, x_3) \in \mathbb{R}^3$ ($x_{\nu} \in \mathbb{R}, \nu = 1, 2, 3$):

$$\begin{aligned} (x_1, x_2, x_3) &= (x_1, 0, 0) + (0, x_2, 0) + (0, 0, x_3) \\ &= x_1(1, 0, 0) + x_2(0, 1, 0) + x_3(0, 0, 1) \\ &= x_1e_1 + x_2e_2 + x_3e_3. \end{aligned}$$

Also ist (e_1, e_2, e_3) ein Erzeugendensystem von \mathbb{R}^3

$$\mathbb{R}^3 = \langle e_1, e_2, e_3 \rangle.$$

b) Seien $u, v \in \mathbb{R}^3$, $u, v \neq 0$ und $u \notin \mathbb{R}v$. Dann ist $\langle u, v \rangle = \{au + bv \mid a, b \in \mathbb{R}\}$ die von u und v aufgespannte Ebene in \mathbb{R}^3 . Ist $c \in \mathbb{R}^3$, so ist

$$\langle u, v \rangle + c := \{au + bv + c \mid a, b \in \mathbb{R}\}$$

ebenfalls eine Ebene (die Ebene durch c mit Richtungsvektoren u, v). Diese Ebene ist nur dann ein Untervektorraum von \mathbb{R}^3 , wenn $0 \in \langle u, v \rangle + c$ gilt, d.h. wenn $c \in \langle u, v \rangle$ ist.

Definition 7.14 Sind W_1, W_2 Untervektorräume eines K -Vektorraumes V , so heißt der Untervektorraum

$$W_1 + W_2 := \text{Span}(W_1 \cup W_2) = \{x + y \mid x \in W_1, y \in W_2\}$$

die Summe von W_1 und W_2 .

Gilt $W_1 \cap W_2 = \{0\}$, so heißt die Summe $W_1 + W_2$ **direkte Summe** von W_1 und W_2 und wird auch mit $W_1 \oplus W_2$ bezeichnet.

Allgemeiner: Für Untervektorräume $W_1, \dots, W_n \subset V$ wird die Summe $W_1 + \dots + W_n$ als der kleinste Untervektorraum von V definiert, der W_1, \dots, W_n enthält:

$$W_1 + \dots + W_n = \text{Span}(W_1 \cup \dots \cup W_n).$$

Man sieht sofort, dass

$$W_1 + \dots + W_n = \{x_1 + \dots + x_n \mid x_i \in W_i \text{ für } i = 1, \dots, n\}$$

gilt. Man muss dazu ja nur zeigen, dass

$$W = \{x \in V \mid \exists x_1 \in W_1, \dots, x_n \in W_n : x = x_1 + \dots + x_n\}$$

ein Untervektorraum von V ist. Zum Beweis dafür seien $x, y \in W$ und $a, b \in K$. Dann gibt es Vektoren $x_i \in W_i, y_i \in W_i$ für $i = 1, \dots, n$, so dass

$$x = x_1 + \dots + x_n \text{ und } y = y_1 + \dots + y_n.$$

Es folgt

$$ax + by = (ax_1 + by_1) + \dots + (ax_n + by_n).$$

Da W_i ein Untervektorraum von V ist, ist $ax_i + by_i \in W_i$ und, somit ist $ax + by \in W$.
□

Man nennt die Summe $W_1 + \dots + W_n$ **direkt** und schreibt dann $W_1 \oplus \dots \oplus W_n$, wenn

$$W_i \cap (W_1 + \dots + W_{i-1} + W_{i+1} + \dots + W_n) = \{0\}$$

für $i = 1, \dots, n$ gilt.

Man schreibt auch

$$\sum_{\nu=1}^n W_\nu = W_1 + \dots + W_n,$$

$$\bigoplus_{\nu=1}^n W_\nu = W_1 \oplus \dots \oplus W_n.$$

Satz 7.15 Es sei V ein K -Vektorraum, und W_1, \dots, W_n seien Untervektorräume von V . Dann gilt für $W = W_1 + \dots + W_n$:

W ist direkte Summe von $W_1, \dots, W_n \iff \forall x \in W$ existieren eindeutig bestimmte $x_1 \in W_1, \dots, x_n \in W_n$, so dass

$$x = x_1 + \dots + x_n.$$

Beweis: “ \Rightarrow ” Es sei $W = W_1 \oplus \dots \oplus W_n$. Es sei $x \in W$ und

$$x = x_1 + \dots + x_n = x'_1 + \dots + x'_n$$

mit $x_i, x'_i \in W_i$ für $i = 1, \dots, n$. Dann ist

$$0 = x''_1 + \dots + x''_n \quad \text{mit } x''_i = x_i - x'_i \in W_i,$$

und somit

$$-x''_i = x''_1 + \dots + x''_{i-1} + x''_{i+1} + \dots + x''_n. \quad (*)$$

Da $-x''_i \in W_i$ und $x''_1 + \dots + x''_{i-1} + x''_{i+1} + \dots + x''_n \in \sum_{\substack{\nu=1 \\ \nu \neq i}}^n W_\nu$, und $W_i \cap \sum_{\substack{\nu=1 \\ \nu \neq i}}^n W_\nu = \{0\}$

gilt, folgt aus (*), dass $x''_i = 0$ gilt, also

$$x_i = x'_i \text{ für } i = 1, \dots, n.$$

“ \Leftarrow ” Ist für ein $i \in \{1, \dots, n\}$ der Durchschnitt

$$W_i \cap \sum_{\substack{\nu=1 \\ \nu \neq i}}^n W_\nu$$

vom Nullvektorraum verschieden, so gibt es ein $x'_i \in W_i \setminus \{0\}$, das auch Element in $W_1 + \dots + W_{i-1} + W_{i+1} + \dots + W_n$ ist, also $x'_i = x_1 + \dots + x_{i-1} + x_{i+1} + \dots + x_n$ für geeignete $x_\nu \in W_\nu$, $\nu = 1, \dots, n$, $\nu \neq i$. Damit haben wir aber eine nicht triviale Darstellung des Nullvektors

$$0 = x_1 + \dots + x_n$$

mit $x_\nu \in W_\nu$ für $\nu = 1, \dots, n$ und $x_i := -x'_i \neq 0$. Diese Darstellung ist von der Darstellung

$$0 = 0 + \dots + 0$$

verschieden. □

Beispiel 7.16 a) Es sei $1 \leq k \leq n$. Dann sind

$$W_1 = \{(x_1, \dots, x_k, 0, \dots, 0) \in \mathbb{R}^n \mid x_i \in \mathbb{R}, i = 1, \dots, k\}$$

und

$$W_2 = \{(0, \dots, 0, x_{k+1}, \dots, x_n) \in \mathbb{R}^n \mid x_i \in \mathbb{R}, i = k+1, \dots, n\}$$

Untervektorräume mit $W_1 \cap W_2 = \{0\}$ und $W_1 + W_2 = \mathbb{R}^n$, also gilt:

$$\mathbb{R}^n = W_1 \oplus W_2.$$

b) Es sei $e_i = (0, \dots, 1, \dots, 0) \in \mathbb{R}^n$ (Die Eins steht an der i -ten Stelle.). Dann gilt:

$$\mathbb{R}^n = \mathbb{R}e_1 \oplus \dots \oplus \mathbb{R}e_n,$$

denn $\mathbb{R}^n = \sum_{\nu=1}^n \mathbb{R}e_\nu$ und $\mathbb{R}e_i \cap \left(\sum_{\substack{\nu=1 \\ \nu \neq i}}^n \mathbb{R}e_\nu \right) = \{0\}$.

Übungen

1. Welche der folgenden Teilmengen von $V = \text{Abb}(\mathbb{R}, \mathbb{R})$ sind Untervektorräume?

(a) $\{f \in V \mid f(0) = f(1) = 0\}$.

(b) $\{f \in V \mid f(0) = f(1) = 5\}$.

(c) $\{f \in V \mid \exists a, b, c \in \mathbb{R} : f(x) = ax^2 + bx + c \text{ für alle } x \in \mathbb{R}\}$.

2. Seien W, W_1, W_2 Untervektorräume eines K -Vektorraumes V . Es gelte $W \subset W_1 + W_2$. Gilt dann $W = (W \cap W_1) + (W \cap W_2)$? Was kann man sagen, wenn $W_1 \subset W$ gilt?

3. Wie viele Untervektorräume hat der Vektorraum \mathbb{F}_3^2 ?

8 Basis und Dimension

Es sei K ein Körper und V ein K -Vektorraum.

Definition 8.1 Es seien $x_1, \dots, x_n \in V$. x_1, \dots, x_n heißen **linear unabhängig** (über K) ((x_1, \dots, x_n) heißt linear unabhängiges System von Vektoren), wenn gilt: Sind $a_1, \dots, a_n \in K$ und gilt $a_1x_1 + \dots + a_nx_n = 0$, so gilt:

$$a_1 = a_2 = \dots = a_n = 0.$$

x_1, \dots, x_n heißen **linear abhängig** (über K) (das n -Tupel (x_1, \dots, x_n) von Vektoren heißt linear abhängig), wenn es Elemente $a_1, \dots, a_n \in K$ gibt, so dass $(a_1, \dots, a_n) \neq 0$ und $a_1x_1 + \dots + a_nx_n = 0$ gilt.

Lemma 8.2 x_1, \dots, x_n sind genau dann linear abhängig, wenn es ein $i \in \{1, \dots, n\}$ gibt, so dass gilt:

$$x_i \in \langle x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n \rangle.$$

Beweis:

- a) Es seien x_1, \dots, x_n linear abhängig. Dann gibt es ein n -Tupel $(a_1, \dots, a_n) \in K^n \setminus \{0\}$ mit

$$a_1x_1 + \dots + a_nx_n = 0.$$

Es sei etwa $a_i \neq 0$. Dann folgt

$$x_i = -\frac{1}{a_i}(a_1x_1 + \dots + a_{i-1}x_{i-1} + a_{i+1}x_{i+1} + \dots + a_nx_n),$$

also

$$x_i \in \langle x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n \rangle.$$

- b) Ist umgekehrt $x_i = \sum_{\substack{\nu=1 \\ \nu \neq i}}^n a_\nu x_\nu$, so folgt $a_1x_1 + \dots + a_nx_n = 0$, wenn man $a_i = -1$ setzt. □

Definition 8.3 Ein n -Tupel (x_1, \dots, x_n) von Vektoren $x_1, \dots, x_n \in V$ heißt eine **Basis** von V (auch **K -Basis**), wenn gilt:

- a) $V = \langle x_1, \dots, x_n \rangle$, d.h. (x_1, \dots, x_n) ist ein Erzeugendensystem von V .
 b) x_1, \dots, x_n sind linear unabhängig.

Lemma 8.4 Ist (x_1, \dots, x_n) eine Basis von V , so gibt es zu jedem Vektor $x \in V$ genau ein n -Tupel $(a_1, \dots, a_n) \in K^n$ von Elementen $a_1, \dots, a_n \in K$, so dass

$$x = a_1x_1 + \dots + a_nx_n$$

gilt. a_i heißt dann i -te Koordinate von x bezüglich der Basis (x_1, \dots, x_n) .

Beweis: Da $V = \langle x_1, \dots, x_n \rangle$ gilt, besitzt jeder Vektor $x \in V$ mindestens eine Darstellung

$$x = a_1x_1 + \dots + a_nx_n \quad (a_i \in K).$$

Ist nun $x = b_1x_1 + \dots + b_nx_n$ eine weitere Darstellung ($b_i \in K$), so gilt mit $c_i := a_i - b_i$ die Gleichung:

$$0 = c_1x_1 + \dots + c_nx_n.$$

Da x_1, \dots, x_n linear unabhängig sind, folgt $c_i = 0$, also

$$a_1 = b_1, \dots, a_n = b_n.$$

□

Bemerkung 8.5 Ist $M \subset \mathbb{N}$ nach oben beschränkt, so gibt es ein eindeutig bestimmtes größtes Element n in M ; es wird mit $\max(M)$ bezeichnet. Ist $M \subset \mathbb{N}$ unbeschränkt, so setzen wir $\max(M) = \infty$.

Definition 8.6 $\dim V = \max\{n \in \mathbb{N} \mid \exists \text{ linear unabhängige Vektoren } x_1, \dots, x_n \in V\}$ heißt die **Dimension** von V . Ist $\dim V \neq \infty$, so heißt V **endlichdimensional**.

Satz 8.7 Ist V endlichdimensional, $\dim V = n$, und sind $x_1, \dots, x_n \in V$ linear abhängig, so ist (x_1, \dots, x_n) eine Basis von V .

Insbesondere gilt: Jeder endlichdimensionale Vektorraum ist auch endlich erzeugt.

Beweis: Wir brauchen nur $V \subset \langle x_1, \dots, x_n \rangle$ zu beweisen. Sei also $x \in V$. Da $\dim V = n$, sind $n + 1$ Vektoren stets linear unabhängig. Dies gilt auch für x_1, \dots, x_n, x . Man hat also eine Gleichung

$$a_1x_1 + \dots + a_nx_n + a_{n+1}x = 0,$$

in der nicht alle a_i gleich Null sind. Es muss dann $a_{n+1} \neq 0$ gelten, weil x_1, \dots, x_n linear unabhängig sind. Es folgt

$$x = -\frac{1}{a_{n+1}}(a_1x_1 + \dots + a_nx_n).$$

□

Satz 8.8 (Basisergänzungssatz): Es seien x_1, \dots, x_r linear unabhängige Vektoren in V . y_1, \dots, y_s seien weitere Vektoren in V , so dass

$$V = \langle x_1, \dots, x_r, y_1, \dots, y_s \rangle.$$

Dann gibt es ein k mit $0 \leq k \leq s$ und Indizes $1 \leq i_1 < \dots < i_k \leq s$, so dass

$$(x_1, \dots, x_r, y_{i_1}, \dots, y_{i_k})$$

eine Basis von V ist.

Beweis: Wähle $1 \leq i_1 < \dots < i_k \leq s$, so dass

$$V = \langle x_1, \dots, x_r, y_{i_1}, \dots, y_{i_k} \rangle,$$

aber

$$V \neq \langle x_1, \dots, x_r, y_{i_1}, \dots, y_{i_{m-1}}, y_{i_{m+1}}, \dots, y_{i_k} \rangle$$

für $m = 1, \dots, k$.

Nach Ummumerieren der y_i dürfen wir $i_\nu = \nu$ annehmen. Es gilt also:

$$V = \langle x_1, \dots, x_r, y_1, \dots, y_k \rangle$$

und

$$V \neq \langle x_1, \dots, x_r, y_1, \dots, y_{m-1}, y_{m+1}, \dots, y_k \rangle \quad (*)$$

für $m = 1, \dots, k$.

Wir zeigen, dass $(x_1, \dots, x_r, y_1, \dots, y_k)$ eine Basis von V ist. Dazu müssen wir nur noch die lineare Unabhängigkeit nachweisen.

Annahme: $\exists (a_1, \dots, a_r, b_1, \dots, b_k) \in K^{r+k} \setminus \{0\}$ mit

$$a_1 x_1 + \dots + a_r x_r + b_1 y_1 + \dots + b_k y_k = 0.$$

Dann können nicht alle b_i gleich Null sein, denn wäre das der Fall, so wären auch alle a_i gleich Null, weil x_1, \dots, x_r nach Voraussetzung linear unabhängig sind. Sei etwa $b_k \neq 0$. Dann folgt

$$y_k \in \langle x_1, \dots, x_r, y_1, \dots, y_{k-1} \rangle,$$

und somit ist

$$V = \langle x_1, \dots, x_r, y_1, \dots, y_{k-1} \rangle.$$

Das steht im Widerspruch zu (*). □

Korollar 8.9 Ist V endlich erzeugt, so besitzt V eine Basis.

Beweis: Man wende Satz 8.8 auf $r = 0$ und ein Erzeugendensystem (y_1, \dots, y_s) von V an. □

Lemma 8.10 Es sei (x_1, \dots, x_r) eine Basis von V , und es seien $y_1, y_2 \in V$. Dann sind

$$x_1, \dots, x_{r-1}, y_1, y_2$$

linear abhängig.

Beweis: Da (x_1, \dots, x_r) eine Basis von V ist, gibt es Elemente $a_i, b_i \in K$, $i = 1, \dots, r$, so dass

$$\begin{aligned} y_1 &= a_1 x_1 + \dots + a_r x_r \\ y_2 &= b_1 x_1 + \dots + b_r x_r. \end{aligned}$$

Ist $a_r = 0$, so zeigt die Gleichung

$$a_1 x_1 + \dots + a_{r-1} x_{r-1} - y_1 = 0,$$

dass x_1, \dots, x_{r-1}, y_1 und damit erst recht $x_1, \dots, x_{r-1}, y_1, y_2$ linear abhängig sind. Ist $a_r \neq 0$, so erhält man aus den beiden Gleichungen durch Eliminieren von x_r die Gleichung

$$y_2 - \frac{b_r}{a_r} y_1 = \left(b_1 - a_1 \frac{b_r}{a_r} \right) x_1 + \dots + \left(b_{r-1} - a_{r-1} \frac{b_r}{a_r} \right) x_{r-1},$$

und somit die lineare Abhängigkeit von

$$x_1, \dots, x_{r-1}, y_1, y_2.$$

Satz 8.11 V sei endlichdimensional.

- Es sei (x_1, \dots, x_r) eine Basis von V und (y_1, \dots, y_s) ein Erzeugendensystem von V . Dann gilt $r \leq s$.
- Ist (x_1, \dots, x_r) eine Basis von V , so gilt $r = \dim V$, d.h.: Alle Basen von V haben dieselbe Länge.
- Man hat auch

$$\dim(V) = \min\{n \in \mathbb{N} \mid \exists x_1, \dots, x_n \in V \text{ mit } V = \langle x_1, \dots, x_n \rangle\}.$$

Beweis: Offensichtlich folgen (b) und (c) aus (a).

zu (a): Sei (x_1, \dots, x_r) eine Basis von V . Dann ist $x_r \notin \langle x_1, \dots, x_{r-1} \rangle$, also (x_1, \dots, x_{r-1}) linear unabhängig, aber keine Basis von V . Nach Satz 8.8 gibt es ein k mit $0 \leq k \leq s$, so dass (evtl. nach Ummumerieren von y_1, \dots, y_s)

$$(x_1, \dots, x_{r-1}, y_1, \dots, y_k)$$

eine Basis von V ist. Es muss $k \geq 1$ sein. Nach Lemma 8.10 muss aber $k = 1$ gelten!

$$(x_1, \dots, x_{r-1}, y_1)$$

ist eine Basis von V . Nun wiederhole man diesen Schluss mit $(x_1, \dots, x_{r-2}, y_1)$, falls $r - 1 \geq 1$ ist. Man erhält eine Basis

$$(x_1, \dots, x_{r-2}, y_1, y_2),$$

wobei man evtl. wieder $\{y_2, \dots, y_s\}$ umnumerierte. Man fährt so fort, bis man schließlich eine Basis

$$(y_1, \dots, y_r)$$

von V erhält. Es muss daher $r \leq s$ gelten. \square

Beispiel 8.12 (a) $\dim K^n = n$. Es sei $e_1 = (1, 0, \dots, 0)$, $e_2 = (0, 1, 0, \dots, 0), \dots$, $e_n = (0, \dots, 0, 1)$. Dann ist (e_1, \dots, e_n) eine Basis von K^n . Diese Basis heißt die **Standardbasis** von K^n .

(b) Eine Funktion $f : K \rightarrow K$ heißt **Polynomfunktion** auf $K : \Leftrightarrow \exists n \in \mathbb{N}$, $a_0, \dots, a_n \in K$, so dass

$$f(x) = \sum_{\nu=0}^n a_\nu x^\nu = a_n x^n + \dots + a_1 x + a_0$$

für alle $x \in K$.

W sei die Menge aller Polynomfunktionen $f : K \rightarrow K$. Die Nullfunktion $f = 0$ ist Element von W .

Sind $a, b \in K$ und $f, g \in W$, so ist die Funktion $af + bg : K \rightarrow K$ auch eine Polynomfunktion. Sei nämlich

$$f(x) = \sum_{\nu=0}^n a_\nu x^\nu, \quad g(x) = \sum_{\nu=0}^m b_\nu x^\nu.$$

Ist $n > m$, so setze man $b_\nu := 0$ für $\nu = m + 1, \dots, n$. Dann ist

$$(af + bg)(x) = \sum_{\nu=0}^n (aa_\nu + bb_\nu)x^\nu.$$

W ist somit ein Untervektorraum von $\text{Abb}(K, K)$.

Wir wollen die Dimension von W bestimmen:

1. Fall: K ist endlich, $|K| = q$. Dann seien a_0, \dots, a_{q-1} die sämtlichen Elemente von K . Dann ist eine Funktion $f : K \rightarrow K$ durch ein q -Tupel $(b_0, \dots, b_{q-1}) \in K^q$ gegeben, indem man

$$f(a_i) := b_i \text{ für } i = 0, \dots, q-1$$

setzt.

$$\text{Abb}(K, K) \rightarrow K^q, \quad f \mapsto (f(a_0), \dots, f(a_{q-1}))$$

ist eine Bijektion. Es sei

$$\delta_i : K \rightarrow K$$

die Funktion

$$\delta_i(a_j) = \begin{cases} 1 & \text{falls } i = j \\ 0 & \text{falls } i \neq j \end{cases}.$$

Dann ist $(\delta_0, \dots, \delta_{q-1})$ eine Basis von $\text{Abb}(K, K)$, denn für $f \in \text{Abb}(K, K)$ gilt

$$f = \sum_{i=0}^{q-1} f(a_i)\delta_i, \text{ also ist}$$

$$\dim \text{Abb}(K, K) = q \neq \infty.$$

Wir sehen weiter, dass in diesem Fall $W = \text{Abb}(K, K)$ gilt, denn es gilt

$$\delta_i(x) = \frac{(x - a_0)(x - a_1) \dots (x - a_{i-1})(x - a_{i+1}) \dots (x - a_{q-1})}{(a_i - a_0)(a_i - a_1) \dots (a_i - a_{i-1})(a_i - a_{i+1}) \dots (a_i - a_{q-1})}$$

und somit ist δ_i für $i = 0, \dots, q - 1$ eine Polynomfunktion auf K . □

2. Fall: K ist unendlich (z.B. $K = \mathbb{Q}, \mathbb{R}, \mathbb{C}$). $f_n \in W$ sei die Polynomfunktion

$$f_n(x) = x^n \text{ für } x \in K \quad (n \in \mathbb{N}).$$

(Dabei ist $x^0 = 1$, $x^n = x \cdot x^{n-1}$ für $n \geq 1$.)

Es gilt nach Definition von W :

$$W = \text{Span} (\{f_0, f_1, f_2, \dots\}).$$

Wir zeigen: $\dim W = \infty$.

Da $\dim W = \max\{n \mid \exists n \text{ linear unabhängige Polynomfunktionen}\}$, genügt es zu zeigen, dass f_0, \dots, f_n für jedes $n \in \mathbb{N}$ linear unabhängig sind. Seien dazu $a_0, \dots, a_n \in K$, und es gelte

$$a_0 f_0 + \dots + a_n f_n = 0.$$

Dann gilt also:

$$a_0 + a_1 x + \dots + a_n x^n = 0 \text{ für alle } x \in K.$$

Eine Polynomfunktion $f(x) = a_n x^n + \dots + a_1 x + a_0$ mit $a_n \neq 0$ hat aber höchstens n verschiedene Nullstellen in K , wie man leicht durch Induktion nach n ($n \geq 0$) beweist. (Der Beweis sei hier eingeschoben: Für $n = 0$ ist die Behauptung offensichtlich richtig. Induktionsschluss $n - 1 \Rightarrow n$: Ist b eine Nullstelle von $f(x) = a_n x^n + \dots + a_1 x + a_0$, so gilt $f(x) = f(x) - f(b) = a_n(x^n - b^n) + \dots + a_1(x - b) = (a_n(x^{n-1} + x^{n-2}b + \dots + b^{n-1}) + \dots + a_1)(x - b) = g(x)(x - b)$. Nach Induktionsvoraussetzung hat g höchstens $n - 1$ Nullstellen, also f höchstens n .) Also gilt: $a_0 = \dots = a_n = 0$. □

Satz 8.13 Es sei V ein K -Vektorraum, und es seien W_1, \dots, W_m endlichdimensionale Untervektorräume, so dass die Summe $W = W_1 + \dots + W_m$ direkt ist. Dann gilt:

$$\dim W = \dim W_1 + \dots + \dim W_m.$$

Beweis: Es sei $n_i = \dim W_i$ und $(x_1^{(i)}, \dots, x_{n_i}^{(i)})$ eine Basis von W_i . Dann ist offensichtlich

$$(x_1^{(1)}, \dots, x_{n_1}^{(1)}, x_1^{(2)}, \dots, x_{n_2}^{(2)}, \dots, x_1^{(m)}, \dots, x_{n_m}^{(m)})$$

ein Erzeugendensystem von W .

Wir müssen noch zeigen, dass dieses System linear unabhängig ist. Sei also

$$a_1^{(1)} x_1^{(1)} + \dots + a_{n_m}^{(m)} x_{n_m}^{(m)} = 0$$

mit $a_j^{(i)} \in K$, $i = 1, \dots, m$, $j = 1, \dots, n_i$.

Dann ist $x_i := a_1^{(i)} x_1^{(i)} + \dots + a_{n_i}^{(i)} x_{n_i}^{(i)} \in W_i$ und $x_1 + \dots + x_m = 0$. Da die Summe direkt ist, folgt (Satz 2.14) $x_1 = \dots = x_m = 0$. Da $x_1^{(i)}, \dots, x_{n_i}^{(i)}$ linear unabhängig sind, folgt aus $x_i = 0$ sofort $a_1^{(i)} = \dots = a_{n_i}^{(i)} = 0$. Dies gilt für $i = 1, \dots, m$. Damit ist der Satz bewiesen. □

Satz 8.14 Es seien W_1, W_2 endlichdimensionale Untervektorräume von V . Dann gilt:

$$\dim W_1 + \dim W_2 = \dim(W_1 + W_2) + \dim(W_1 \cap W_2).$$

Beweis: Es sei (x_1, \dots, x_n) eine Basis von $W_1 \cap W_2$ und

$$\begin{aligned} (x_1, \dots, x_n, y_1, \dots, y_k) & \text{ Basis von } W_1, \\ (x_1, \dots, x_n, y'_1, \dots, y'_l) & \text{ Basis von } W_2. \end{aligned}$$

Wir zeigen, dass dann

$$\mathcal{B} = (x_1, \dots, x_n, y_1, \dots, y_k, y'_1, \dots, y'_l)$$

eine Basis von $W_1 + W_2$ ist.

(a) \mathcal{B} ist ein Erzeugendensystem von $W_1 + W_2$: Sei dazu $x \in W_1 + W_2$. Dann ist $x = x' + x''$ für geeignete Vektoren $x' \in W_1$, $x'' \in W_2$. Es sei

$$\begin{aligned} x' &= a_1x_1 + \dots + a_nx_n + b_1y_1 + \dots + b_ky_k \\ x'' &= a'_1x_1 + \dots + a'_nx_n + b'_1y'_1 + \dots + b'_ly'_l \end{aligned}$$

mit $a_i, a'_i \in K$ für $i = 1, \dots, n$; $b_i \in K$ für $i = 1, \dots, k$ und $b'_i \in K$ für $i = 1, \dots, l$. Dann ist

$$x = (a_1 + a'_1)x_1 + \dots + (a_n + a'_n)x_n + b_1y_1 + \dots + b_ky_k + b'_1y'_1 + \dots + b'_ly'_l,$$

also $x \in \text{Span}(\mathcal{B})$.

(b) \mathcal{B} ist ein linear unabhängiges System: Es seien $a_1, \dots, a_n, b_1, \dots, b_k, b'_1, \dots, b'_l \in K$, und es sei

$$a_1x_1 + \dots + a_nx_n + b_1y_1 + \dots + b_ky_k + b'_1y'_1 + \dots + b'_ly'_l = 0.$$

Wir müssen zeigen, dass alle a_i, b_i, b'_i verschwinden. Es sei

$$\begin{aligned} x' &= a_1x_1 + \dots + a_nx_n + b_1y_1 + \dots + b_ky_k \\ x'' &= b'_1y'_1 + \dots + b'_ly'_l. \end{aligned}$$

Dann gilt $x' \in W_1$, $x'' \in W_2$ und $x' = -x''$, also $x' \in W_1 \cap W_2$. Da (x_1, \dots, x_n) eine Basis von $W_1 \cap W_2$ und $(x_1, \dots, x_n, y_1, \dots, y_k)$ eine Basis von W_1 ist, folgt dann sofort $b_1 = \dots = b_k = 0$.

Da $(x_1, \dots, x_n, y'_1, \dots, y'_l)$ linear unabhängiges System ist, folgt weiter: $a_1 = \dots = a_n = b'_1 = \dots = b'_l = 0$. \square

Zum Schluss wollen wir untersuchen, wie man Basen von Untervektorräumen von K^n bestimmen kann. Dazu beweisen wir zunächst

Lemma 8.15 Es sei V ein K -Vektorraum, und es seien $x_1, \dots, x_m \in V$. Dann gilt

$$(1) \quad \langle x_1, \dots, x_i, \dots, x_j, \dots, x_n \rangle = \langle x_1, \dots, x_j, \dots, x_i, \dots, x_m \rangle$$

für alle $1 \leq i < j \leq m$.

(2)

$$\langle x_1, \dots, x_j, \dots, x_i, \dots, x_m \rangle = \langle x_1, \dots, x_j, \dots, x_i + ax_j, \dots, x_m \rangle$$

für alle $i \neq j$ und $a \in K$. x_i wird durch $x_i + ax_j$ ersetzt.

(3)

$$\langle x_1, \dots, x_i, \dots, x_m \rangle = \langle x_1, \dots, ax_i, \dots, x_m \rangle$$

für alle $1 \leq i \leq m$ und $a \in K$, $a \neq 0$.

(4)

$$\langle x_1, \dots, x_m \rangle = \langle x_1, \dots, x_{m-1} \rangle,$$

falls $x_m = 0$.**Beweis:** (1) ist trivial.zu (2): Ist $x = \sum_{\nu=1}^m a_\nu x_\nu$, so ist

$$x = a_1 x_1 + \dots + a_{j-1} x_{j-1} + (a_j - aa_j) x_j + \dots + a_i (x_i + ax_j) + \dots + a_m x_m,$$

also gilt:

$$\langle x_1, \dots, x_n \rangle \subset \langle x_1, \dots, x_j, \dots, x_i + ax_j, \dots, x_m \rangle.$$

Offensichtlich ist

$$\langle x_1, \dots, x_j, \dots, x_i + ax_j, \dots, x_m \rangle \subset \langle x_1, \dots, x_j, \dots, x_i, \dots, x_m \rangle.$$

(3) und (4) sind trivial. □Es seien nun $a_1, \dots, a_m \in K^n$, und $W = \langle a_1, \dots, a_m \rangle$ sei der von a_1, \dots, a_m aufgespannte Untervektorraum von K^n .Lemma 8.15 ermöglicht es nun, eine Basis von W und damit auch $\dim W$ zu bestimmen. Wir treffen wieder auf das uns schon vertraute Gaußsche Eliminationsverfahren:

Es sei

$$a_i = (a_{i1}, \dots, a_{in}) \text{ für } i = 1, \dots, m.$$

Man erhält dann eine Matrix

$$A = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix} = \begin{pmatrix} a_1 \\ \vdots \\ a_m \end{pmatrix},$$

deren Zeilen die Vektoren $a_1, \dots, a_m \in K^n$ sind, die also den Vektorraum W aufspannen. Nach Lemma 8.15 darf man nun die folgenden elementaren Zeilenumformungen an der Matrix A durchführen, ohne den von den Zeilen aufgespannten Untervektorraum, also ohne W zu ändern.

Umformungen vom Typ I:

 T_{ij} : Vertausche die i -te und j -te Zeile.

Umformungen vom Typ II:

$T_{ij}(c)$: Ersetze die i -te Zeile durch die Zeile $a_i + ca_j$

Hierbei sind i, j Indizes mit $1 \leq i \leq m; 1 \leq j \leq m; i \neq j$ und c ist ein Element aus K .

Umformungen vom Typ III:

$T_i(c)$: Ersetze die i -te Zeile durch ihr c -faches.

Hierbei ist c ein von Null verschiedenes Element in K .

Umformung vom Typ IV:

T_i : Weglassen der i -ten Zeile, falls diese Null ist.

Man geht nun so vor (Gaußsches Eliminationsverfahren):

1. Schritt: Man sucht die erste von Null verschiedene Spalte von A . Sei dies die j_1 -te Spalte. Dann suche man einen Index i mit $a_{ij_1} \neq 0$. Ist $i \neq 1$, so wende man die elementare Zeilenumformung T_{1i} an. Die neue Matrix nennen wir wieder A . Jetzt gilt $a_{1j_1} \neq 0$. Für $i = 2, \dots, m$ wende man nun $T_{i,1} \left(-\frac{a_{ij_1}}{a_{1j_1}} \right)$ an. Dann hat A die Gestalt

$$A = \begin{pmatrix} 0 & \dots & 0 & a_{1j_1} & * & \dots & * \\ 0 & & 0 & 0 & \boxed{A'} & & \\ \vdots & & \vdots & \vdots & & & \\ 0 & & 0 & 0 & & & \end{pmatrix}.$$

2. Schritt: Mit der $(m-1) \times (n-j_1)$ -Matrix A' wiederholt man das Verfahren und erhält

$$A = \begin{pmatrix} 0 & \dots & 0 & a_{1j_1} & * & \dots & * & * & * & \dots & * \\ 0 & \dots & 0 & 0 & 0 & \dots & 0 & a_{2j_2} & * & \dots & * \\ 0 & & \dots & & & & 0 & \boxed{A''} & & & \\ \vdots & & & & & & \vdots & & & & \\ 0 & & \dots & & & & 0 & & & & \end{pmatrix}.$$

Nach endlich vielen Schritten hat man eine *Zeilenstufenform* erhalten, d.h. aus A erhält man durch elementare Zeilenumformungen eine Matrix B der Form

$$B = \begin{pmatrix} & b_{1j_1} & & & & \\ & & b_{2j_2} & & * & \\ 0 & & & b_{3j_3} & & \\ & & & & \ddots & \\ & & & & & b_{rj_r} \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \\ \vdots \\ b_r \end{pmatrix}$$

mit $1 \leq j_1 < \dots < j_r \leq n$ und $b_{1j_1} \neq 0, \dots, b_{rj_r} \neq 0$. Durch Anwenden von $T_i(b_{ij_i}^{-1})$ kann man sogar erreichen, dass $b_{ij_i} = 1$ für $i = 1, \dots, r$, bzw. dass eine reduzierte Zeilenstufenform vorliegt.

Nach Lemma 8.4 gilt dann:

$$W = \langle a_1, \dots, a_m \rangle = \langle b_1, \dots, b_r \rangle.$$

Man sieht leicht, dass b_1, \dots, b_r linear unabhängig sind: Seien $\lambda_1, \dots, \lambda_r \in K$, und gelte

$$\lambda_1 b_1 + \dots + \lambda_r b_r = 0;$$

dann gilt:

$$\lambda_1 b_{1i} + \dots + \lambda_r b_{ri} = 0 \text{ für } i = 1, \dots, n$$

und somit

$$\begin{aligned} 0 &= \lambda_1 b_{1j_1} \\ 0 &= \lambda_1 b_{1j_2} + \lambda_2 b_{2j_2} \\ 0 &= \lambda_1 b_{1j_3} + \lambda_2 b_{2j_2} + \lambda_3 b_{3j_3} \\ &\vdots \\ 0 &= \lambda_1 b_{1j_r} + \dots + \lambda_r b_{rj_r} \end{aligned}$$

Da $b_{ij_i} \neq 0$ für $i = 1, \dots, r$, folgt daraus unmittelbar $\lambda_1 = \lambda_2 = \dots = \lambda_r = 0$. (b_1, \dots, b_r) ist also eine Basis von W .

Im allgemeinen ist $0 \leq r \leq m$. Ist $r = m$, so ist (a_1, \dots, a_m) ein Erzeugendensystem von W mit $m = \dim W$. (a_1, \dots, a_m) muss dann schon eine Basis sein (Satz 8.8).

Beispiel 8.16

$$W = \text{Span} \{(1, 0, 1, 1, 1, 1), (1, 0, 1, 1, 0, 1), (1, 0, 0, 0, 1, 1), (1, 1, 1, 0, 0, 1)\}$$

ist 4-dimensionaler Unterraum von \mathbb{F}_2^6 , denn

$$\begin{pmatrix} 1 & 0 & 1 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 & 0 & 1 \\ 1 & 0 & 0 & 0 & 1 & 1 \\ 1 & 1 & 1 & 0 & 0 & 1 \end{pmatrix} \rightarrow \begin{pmatrix} 1 & 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 1 & 0 \end{pmatrix} \text{ durch } T_{21}(1), T_{31}(1) \text{ und } T_{41}(1).$$

Durch Zeilenvertauschungen ergibt sich eine Zeilenstufenform

$$\begin{pmatrix} 1 & 0 & 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{pmatrix}.$$

Also ist $\dim W = 4$.

Übungen

1. Die Elemente aus \mathbb{R}^3 werden als Punkte im Anschauungsraum veranschaulicht. Es seien nun $a_1, a_2, a_3 \in \mathbb{R}^3$. Welche geometrische Figur beschreiben die Mengen

$$\begin{aligned} \text{a) } M &= \{\lambda_1 a_1 + \lambda_2 a_2 + \lambda_3 a_3 \mid \lambda_i \in \mathbb{R}, -1 \leq \lambda_i \leq 1, i = 1, 2, 3\}, \\ \text{b) } D &= \{\lambda_1 a_1 + \lambda_2 a_2 + \lambda_3 a_3 \mid \lambda_i \in \mathbb{R}, \lambda_i \geq 0, \lambda_1 + \lambda_2 + \lambda_3 \leq 1\}, \\ \text{c) } T &= \{\lambda_1 a_1 + \lambda_2 a_2 + \lambda_3 a_3 \mid \lambda_i \in \mathbb{R}, \lambda_i \geq 0, \lambda_1 + \lambda_2 + \lambda_3 = 1\}, \\ \text{d) } K &= \{\lambda_1 a_1 + \lambda_2 a_2 + \lambda_3 a_3 \mid \lambda_i \in \mathbb{R}, \lambda_1^2 + \lambda_2^2 + \lambda_3^2 = 1\}. \end{aligned}$$

Es sei vorausgesetzt, dass a_1, a_2, a_3 linear unabhängig sind.

2. Es sei V ein K -Vektorraum, und es sei $n \geq 5$. x_1, \dots, x_n seien linear unabhängige Vektoren in V . Es sei

$$W_1 = \langle x_1 + x_2, x_1 + x_3, x_3 - x_1 \rangle, \quad W_2 = \langle x_1 + 2x_2, x_3 + x_4 \rangle, \quad W_3 = \langle x_4, x_5, \dots, x_n \rangle.$$

Bestimmen Sie eine Basis und die Dimension der Räume $W_1, W_2, W_3, W_1 + W_3, W_1 \cap W_2, W_1 \cap W_3, W_2 \cap W_3$.

3. Es sei V der Vektorraum der Polynomfunktionen $f : \mathbb{R} \rightarrow \mathbb{R}$ vom Grad $\leq n$. Ist

$$W = \{f \in V \mid \forall x \in \mathbb{R} : f(x) = f(-x)\}$$

ein Untervektorraum? Wenn ja, bestimmen Sie eine Basis und die Dimension von W (Beachten Sie: Ist $f \in V$, $f \neq 0$, so gibt es höchstens n verschiedene Zahlen $a \in \mathbb{R}$ mit $f(a) = 0$).

4. Es sei P der Vektorraum der Polynomfunktionen $f : \mathbb{R} \rightarrow \mathbb{R}$. Zeigen Sie:

$$W = \{f \in P \mid f(0) + f(1) + f(2) = 0\}$$

ist ein unendlichdimensionaler Untervektorraum von P .

5. (a) Es sei $k \leq n$, und $x_1, \dots, x_k \in K^n$ seien linear unabhängig. Beschreiben Sie ein konstruktives Verfahren, mit dem man (x_1, \dots, x_k) zu einer Basis (x_1, \dots, x_n) von K^n ergänzen kann.
 (b) Es sei $W = \langle x_1, x_2, x_3 \rangle \subset \mathbb{R}^5$ mit

$$x_1 = (1, 1, 0, 2, 3), \quad x_2 = (1, 2, 3, 4, 0), \quad x_3 = (2, 0, 0, 1, 4).$$

Konstruieren Sie unendlich viele verschiedene Untervektorräume $W' \subset \mathbb{R}^5$ mit $\mathbb{R}^5 = W \oplus W'$.

6. Es sei V ein K -Vektorraum, $x_1, x_2, x_3, x_4 \in V$. Wieso sind die fünf Vektoren

$$\begin{aligned} y_1 &= x_1 + x_2 + x_3 + x_4, \quad y_2 = 2x_1 + 2x_2 + x_3 + x_4, \quad y_3 = x_1 + x_2 + 3x_3 - x_4, \\ y_4 &= x_1 - x_3 + x_4, \quad y_5 = -x_2 + x_3 - x_4 \end{aligned}$$

stets linear abhängig?

7. Es sei V ein endlichdimensionaler \mathbb{C} -Vektorraum. Durch Einschränkung der skalaren Multiplikation $\cdot : \mathbb{C} \times V \rightarrow V$ auf $\mathbb{R} \times V$ wird V ein \mathbb{R} -Vektorraum. Zeigen Sie: Ist (x_1, \dots, x_n) eine Basis von V über \mathbb{C} , so ist $(x_1, \dots, x_n, ix_1, \dots, ix_n)$ eine Basis von V über \mathbb{R} .
8. Es sei $n \geq 0$. Für welche Paare $(r, s) \in \mathbb{N} \times \mathbb{N}$ gibt es einen \mathbb{C} -Vektorraum V , Vektoren $x_1, \dots, x_n \in V$, so dass

$$r = \dim_{\mathbb{R}} \left(\sum_{\nu=1}^n \mathbb{C}x_{\nu} \right), \quad s = \dim_{\mathbb{R}} \left(\sum_{\nu=1}^n \mathbb{R}x_{\nu} \right)$$

gilt (Dabei bedeutet $\dim_{\mathbb{R}}(W)$ die reelle Dimension des Vektorraumes W , also die maximale Anzahl über \mathbb{R} linear unabhängiger Vektoren in W).

9. Bestimmen Sie eine Basis des von den Vektoren $(1, 2, 0, 3)$, $(2, 5, 2, 4)$, $(1, 4, 7, 2)$ aufgespannten Untervektorraumes von \mathbb{F}_{11}^4 .
10. Es seien V_1, V_2 K -Vektorräume. Zeigen Sie, dass $V = V_1 \times V_2$ auf natürliche Weise ein K -Vektorraum wird mit $W_1 = V_1 \times \{0\}$, $W_2 = \{0\} \times V_2$ als Untervektorräumen mit $V = W_1 \oplus W_2$.

Bemerkung 8.17 Wir haben nur Basen in endlichdimensionalen Vektorräumen betrachtet. Allgemein kann man definieren: Eine Teilmenge B eines K -Vektorraumes V heißt *Basis* von V , wenn gilt:

- a) B ist ein Erzeugendensystem von V , d.h.

$$V = \text{Span}(B).$$

- b) Sind $x_1, \dots, x_n \in B$ beliebig, aber paarweise verschieden, so sind x_1, \dots, x_n linear unabhängig.

Mit Hilfe des Zornschen Lemmas kann man dann zeigen, dass jeder K -Vektorraum eine Basis besitzt. Wir gehen darauf nicht ein, zumal in den Gebieten der Mathematik, wo unendlichdimensionale Vektorräume auftreten, dieser Basisbegriff relativ unwichtig ist. Andere Basen, z.B. Hilbertraumbasen von Hilberträumen, sind dann zu betrachten.

Wer dennoch einen Beweis sehen möchte, schaue etwa bei [24] oder [17].

9 Lineare Abbildungen

Es sei K ein Körper.

Definition 9.1 Es seien V, W K -Vektorräume. Eine Abbildung $\varphi : V \rightarrow W$ heißt **linear** (genauer: K -linear), wenn gilt:

$$\varphi(x + y) = \varphi(x) + \varphi(y) \quad \text{für alle } x, y \in V, \quad (28)$$

$$\varphi(ax) = a\varphi(x) \quad \text{für alle } x \in V, a \in K. \quad (29)$$

Man sagt auch: φ ist ein **Homomorphismus** von Vektorräumen oder ein **Vektorraumhomomorphismus**.

Aus der Definition folgt leicht

$$\varphi(0) = 0, \quad (30)$$

denn $\varphi(x) = \varphi(x + 0) = \varphi(x) + \varphi(0)$, sowie

$$\varphi(-x) = -\varphi(x), \quad (31)$$

denn $\varphi(x) + \varphi(-x) = \varphi(x + (-x)) = \varphi(0) = 0$.

Sind $x_1, \dots, x_n \in V$ und $a_1, \dots, a_n \in K$, so folgt durch Induktion aus (28) und (29) auch die Gleichung

$$\varphi \left(\sum_{\nu=1}^n a_\nu x_\nu \right) = \sum_{\nu=1}^n a_\nu \varphi(x_\nu); \quad (32)$$

und weiter

$$x_1, \dots, x_n \text{ linear abhängig} \Rightarrow \varphi(x_1), \dots, \varphi(x_n) \text{ linear abhängig}; \quad (33)$$

$$\varphi(x_1), \dots, \varphi(x_n) \text{ linear unabhängig} \Rightarrow x_1, \dots, x_n \text{ linear unabhängig}. \quad (34)$$

Lemma 9.2 Es seien V, V', V'' K -Vektorräume.

- Die identische Abbildung $id_V : V \rightarrow V$ ist linear.
- Sind $\varphi : V \rightarrow V', \psi : V' \rightarrow V''$ linear, so ist auch $\psi \circ \varphi : V \rightarrow V''$ linear.

Beweis: (a) ist trivial.

zu (b):

$$\begin{aligned} \psi \circ \varphi(x + y) &= \psi(\varphi(x + y)) = \psi(\varphi(x) + \varphi(y)) = \psi(\varphi(x)) + \psi(\varphi(y)), \\ &= \psi \circ \varphi(x) + \psi \circ \varphi(y), \\ \psi \circ \varphi(ax) &= \psi(\varphi(ax)) = \psi(a\varphi(x)) = a\psi(\varphi(x)) = a(\psi \circ \varphi)(x). \end{aligned}$$

□

Lemma 9.3 Sind $\varphi, \psi : V \rightarrow W$ zwei lineare Abbildungen und ist $V = \langle x_1, \dots, x_n \rangle$ und $\varphi(x_i) = \psi(x_i)$ für $i = 1, \dots, n$, so gilt: $\varphi = \psi$;

d.h. eine lineare Abbildung ist schon durch ihre Werte auf einem Erzeugendensystem festgelegt.

Beweis: Sei $x \in V$. Dann gibt es Elemente $a_1, \dots, a_n \in K$, so dass $x = \sum_{\nu=1}^n a_\nu x_\nu$. Also gilt nach (32):

$$\varphi(x) = \sum_{\nu=1}^n a_\nu \varphi(x_\nu) = \sum_{\nu=1}^n a_\nu \psi(x_\nu) = \psi(x).$$

□

Satz 9.4 Es seien V, W K -Vektorräume, und es sei (x_1, \dots, x_n) eine Basis von V . y_1, \dots, y_n seien beliebige Vektoren in W . Dann gibt es genau eine lineare Abbildung $\varphi : V \rightarrow W$ mit $\varphi(x_\nu) = y_\nu$ für $\nu = 1, \dots, n$.

Beweis: Wir konstruieren eine Abbildung $\varphi : V \rightarrow W$. Es sei $x \in V$. Da (x_1, \dots, x_n) eine Basis von V ist, gibt es eindeutig bestimmte Elemente $a_1, \dots, a_n \in K$, so dass

$$x = a_1 x_1 + \dots + a_n x_n$$

gilt. Man setze

$$\varphi(x) := a_1 y_1 + \dots + a_n y_n.$$

Offensichtlich gilt: $\varphi(x_\nu) = y_\nu$ für $\nu = 1, \dots, n$. Wir zeigen nun, dass φ linear ist:

a) Ist $x = \sum_{\nu=1}^n a_\nu x_\nu$, $y = \sum_{\nu=1}^n b_\nu x_\nu$, so gilt:

$$\begin{aligned} \varphi(x+y) &= \varphi\left(\sum_{\nu=1}^n (a_\nu + b_\nu)x_\nu\right) = \sum_{\nu=1}^n (a_\nu + b_\nu)y_\nu \\ &= \sum_{\nu=1}^n a_\nu y_\nu + \sum_{\nu=1}^n b_\nu y_\nu = \varphi(x) + \varphi(y). \end{aligned}$$

b) Ist $x = \sum_{\nu=1}^n a_\nu x_\nu$ und $a \in K$, so ist

$$\begin{aligned} \varphi(ax) &= \varphi\left(\sum_{\nu=1}^n (aa_\nu)x_\nu\right) = \sum_{\nu=1}^n (aa_\nu)y_\nu \\ &= a \sum_{\nu=1}^n a_\nu y_\nu = a\varphi(x). \end{aligned}$$

Damit ist die Existenz von φ bewiesen. Die Eindeutigkeit folgt aus Lemma 9.3. □

Satz 9.4 liefert eine Methode zur Konstruktion linearer Abbildungen eines endlichdimensionalen Vektorraumes V in einen Vektorraum W : Man wähle irgendeine Basis (x_1, \dots, x_n) von V und setze $B = \{x_1, \dots, x_n\} \subset V$. Jede beliebige Abbildung $f : B \rightarrow W$ ($f(x_\nu) = y_\nu$) kann man dann auf eindeutige Weise linear auf ganz V

fortsetzen. Ist $i : B \rightarrow V$ die Inklusionsabbildung, so hat man ein kommutatives Diagramm

$$\begin{array}{ccc} B & \xrightarrow{f} & W \\ \downarrow i & \searrow \varphi & \\ V & & \end{array} \quad (\varphi \circ i = f).$$

φ ist die eindeutig bestimmte lineare Fortsetzung von f .

Definition 9.5 Es seien V, W beliebige K -Vektorräume. Ein **Isomorphismus** $\varphi : V \rightarrow W$ ist eine bijektive lineare Abbildung $\varphi : V \rightarrow W$. V und W heißen **isomorph** ($V \cong W$), wenn es einen Isomorphismus $\varphi : V \rightarrow W$ gibt.

Lemma 9.6 Es seien V, V', V'' K -Vektorräume. Dann gilt:

- (i) id_V ist ein Isomorphismus.
- (ii) Sind $\varphi : V \rightarrow V', \psi : V' \rightarrow V''$ Isomorphismen, so ist auch $\psi \circ \varphi : V \rightarrow V''$ ein Isomorphismus.
- (iii) Ist $\varphi : V \rightarrow V'$ ein Isomorphismus, so ist die Umkehrabbildung $\varphi^{-1} : V' \rightarrow V$ auch ein Isomorphismus.
- (iv) Eine lineare Abbildung $\varphi : V \rightarrow V'$ ist genau dann ein Isomorphismus, wenn es eine lineare Abbildung $\psi : V' \rightarrow V$ gibt mit

$$\varphi \circ \psi = id_{V'}, \quad \psi \circ \varphi = id_V.$$

Beweis: (i), (ii) sind klar.

zu (iii): $\varphi : V \rightarrow V'$ sei linear und bijektiv. Wir zeigen, dass auch $\varphi^{-1} : V' \rightarrow V$ linear ist, bijektiv ist es ohnehin. Seien $x', y' \in V'$ und $x = \varphi^{-1}(x'), y = \varphi^{-1}(y')$. Dann ist also $\varphi(x) = x'$ und $\varphi(y) = y'$ und somit

$$\varphi(x + y) = \varphi(x) + \varphi(y) = x' + y',$$

also

$$\varphi^{-1}(x' + y') = x + y = \varphi^{-1}(x') + \varphi^{-1}(y').$$

Ist $a \in K$, so gilt

$$\varphi(ax) = a\varphi(x) = ax',$$

also

$$\varphi^{-1}(ax') = ax = a\varphi^{-1}(x').$$

Zu (iv): Wegen (iii) ist dies klar. □

Satz 9.7 Es sei V ein endlichdimensionaler K -Vektorraum, $n = \dim V$. Dann gilt:

$$V \cong K^n.$$

Dabei wird $K^0 = \{0\}$ gesetzt.

Beweis: Es sei (x_1, \dots, x_n) eine Basis von V . (e_1, \dots, e_n) sei die Standardbasis von K^n . Nach Satz 9.4 gibt es eindeutig bestimmte lineare Abbildungen $\varphi : V \rightarrow K^n$ mit $\varphi(x_\nu) = e_\nu$ für $\nu = 1, \dots, n$ und $\psi : K^n \rightarrow V$ mit $\psi(e_\nu) = x_\nu$ für $\nu = 1, \dots, n$. Dann sind $\varphi \circ \psi$ und id_{K^n} lineare Abbildungen $K^n \rightarrow K^n$, die e_ν in e_ν abbilden für $\nu = 1, \dots, n$. Nach Lemma 9.3 gilt also $\varphi \circ \psi = id_{K^n}$. Genauso folgt $\psi \circ \varphi = id_V$. Es folgt $V \cong K^n$ nach Lemma 9.6 (iv). \square

Isomorphe Vektorräume haben dieselbe Vektorraumstruktur. Alle Größen oder Eigenschaften eines Vektorraumes, die nur von der Vektorraumstruktur abhängen, bleiben erhalten, wenn man zu einem isomorphen Vektorraum übergeht. Zum Beispiel gilt

Satz 9.8 Sind V, W beliebige K -Vektorräume mit $V \cong W$, so ist $\dim V = \dim W$.

Beweis: Es sei $\varphi : V \rightarrow W$ ein Isomorphismus. Da φ injektiv ist, folgt nun leicht: Sind $x_1, \dots, x_n \in V$ linear unabhängig, so sind $\varphi(x_1), \dots, \varphi(x_n)$ linear unabhängig. Denn aus $a_1\varphi(x_1) + \dots + a_n\varphi(x_n) = 0$ folgt $\varphi(a_1x_1 + \dots + a_nx_n) = 0 = \varphi(0)$, also $a_1x_1 + \dots + a_nx_n = 0$ und somit $a_1 = \dots = a_n = 0$.

Es ergibt sich also $\dim V \leq \dim W$. Derselbe Schluss für φ^{-1} ergibt $\dim W \leq \dim V$. Hierbei haben wir $\infty \leq \infty, n \leq \infty$ für alle $n \in \mathbb{N}$ vereinbart. \square

Korollar 9.9 Sind V, W endlichdimensionale K -Vektorräume, so gilt:

$$V \cong W \quad \Leftrightarrow \quad \dim V = \dim W.$$

Satz 9.10 Es sei $\varphi : V \rightarrow V'$ eine lineare Abbildung.

- a) Ist $W' \subset V'$ ein Untervektorraum, so ist $\varphi^{-1}(W')$ ein Untervektorraum von V .
- b) Ist $W \subset V$ ein Untervektorraum, so ist $\varphi(W)$ ein Untervektorraum von V' .

Beweis: zu a): Da $\varphi(0) = 0 \in W'$, ist $0 \in \varphi^{-1}(W')$. Seien $x, y \in \varphi^{-1}(W')$ und $a, b \in K$, dann gilt: $\varphi(x), \varphi(y) \in W'$ und somit auch

$$\varphi(ax + by) = a\varphi(x) + b\varphi(y) \in W',$$

d.h. $ax + by \in \varphi^{-1}(W')$.

zu b): Da $W \neq \emptyset$ ist, ist auch $\varphi(W) = \{\varphi(x) | x \in W\} \neq \emptyset$. Seien $x', y' \in \varphi(W)$, $a, b \in K$. Es gibt also Vektoren $x, y \in W$ mit $\varphi(x) = x'$ und $\varphi(y) = y'$. Da W Untervektorraum von V ist, ist $ax + by \in W$; und somit

$$ax' + by' = a\varphi(x) + b\varphi(y) = \varphi(ax + by) \in \varphi(W).$$

\square

Definition 9.11 Es sei $\varphi : V \rightarrow W$ eine lineare Abbildung von K -Vektorräumen. Dann heißt der Untervektorraum

$$\ker \varphi := \{x \in V \mid \varphi(x) = 0\} = \varphi^{-1}(\{0\})$$

von V der **Kern** von φ und der Untervektorraum

$$\operatorname{im} \varphi := \varphi(V) = \{y \in W \mid \exists x \in V : y = \varphi(x)\}$$

von W das **Bild** von φ ($\ker = \text{kernel}$, $\operatorname{im} = \text{image}$).

Ein oft benutztes Kriterium für die Injektivität einer linearen Abbildung ist

Lemma 9.12 Es sei $\varphi : V \rightarrow W$ eine lineare Abbildung. Dann gilt:

$$\varphi \text{ ist injektiv} \Leftrightarrow \ker \varphi = \{0\}.$$

(Statt $\ker \varphi = \{0\}$ schreiben wir in Zukunft kurz $\ker \varphi = 0$.)

Beweis: Sei φ injektiv. Ist $x \in \ker \varphi$, so ist $\varphi(x) = 0 = \varphi(0)$, also $x = 0$. Sei umgekehrt $\ker \varphi = 0$. Sind $x, y \in V$ mit $\varphi(x) = \varphi(y)$, so gilt $\varphi(x - y) = \varphi(x) - \varphi(y) = 0$, also $x - y \in \ker \varphi$ und somit $x = y$. φ ist also injektiv. \square

Definition 9.13 Es sei $\varphi : V \rightarrow W$ eine lineare Abbildung.

- a) $\operatorname{rg} \varphi := \dim(\operatorname{im} \varphi)$ heißt der **Rang** von φ .
- b) $\operatorname{crg} \varphi := \dim(\ker \varphi)$ heißt der **Corang** von φ .

Satz 9.14 Es sei $\varphi : V \rightarrow W$ eine lineare Abbildung, und V sei endlichdimensional. Dann sind auch $\ker \varphi$ und $\operatorname{im} \varphi$ endlichdimensional, und es gilt:

$$\dim V = \operatorname{rg} \varphi + \operatorname{crg} \varphi.$$

Beweis: Da $\ker \varphi$ Untervektorraum von V ist, ist $\ker \varphi$ endlichdimensional. Sei (x_1, \dots, x_r) eine Basis von $\ker \varphi$. Man ergänze dieses System (x_1, \dots, x_r) zu einer Basis (x_1, \dots, x_n) von V ($r \leq n$).

Um den Satz zu beweisen, brauchen wir nur noch zu zeigen, dass

$$(\varphi(x_{r+1}), \dots, \varphi(x_n))$$

eine Basis von $\operatorname{im} \varphi$ ist.

Zunächst zeigen wir, dass $\varphi(x_{r+1}), \dots, \varphi(x_n)$ linear unabhängig sind. Seien $a_{r+1}, \dots, a_n \in K$ und gelte:

$$a_{r+1}\varphi(x_{r+1}) + \dots + a_n\varphi(x_n) = 0.$$

Dann ist $\varphi(a_{r+1}x_{r+1} + \dots + a_nx_n) = 0$, also

$$a_{r+1}x_{r+1} + \dots + a_nx_n \in \ker \varphi = \langle x_1, \dots, x_r \rangle;$$

es folgt $a_{r+1} = \dots = a_n = 0$.

Jetzt zeigen wir noch, dass im φ von $\varphi(x_{r+1}), \dots, \varphi(x_n)$ erzeugt wird. Es sei $y \in \text{im } \varphi$, also $y = \varphi(x)$ für ein $x \in V$. Ist $x = \sum_{\nu=1}^n a_\nu x_\nu$, so folgt

$$y = \varphi(x) = \sum_{\nu=r+1}^n a_\nu \varphi(x_\nu) \in \langle \varphi(x_{r+1}), \dots, \varphi(x_n) \rangle,$$

denn $\varphi(x_\nu) = 0$ für $\nu = 1, \dots, r$. □

Es folgen nun einige Beispiele.

Beispiel 9.15

a) $\varphi : K^2 \rightarrow K^2$ ist K -linear $\Leftrightarrow \exists a, b, c, d \in K$, so dass

$$\varphi(x, y) = (ax + by, cx + dy).$$

Beweis: Sei $\varphi(1, 0) = (a, c)$, $\varphi(0, 1) = (b, d)$. Dann ist

$$\varphi(x, y) = x\varphi(1, 0) + y\varphi(0, 1) = (ax + by, cx + dy).$$

Die 2×2 -Matrix

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

heißt die Matrix von φ . Es gilt:

$$\ker \varphi = \{(x, y) \in K^2 \mid ax + by = 0, cx + dy = 0\}.$$

Also gilt $\ker \varphi = 0$ genau dann, wenn das homogene lineare Gleichungssystem

$$\begin{aligned} ax + by &= 0 \\ cx + dy &= 0 \end{aligned}$$

nur die triviale Lösung $(x, y) = (0, 0)$ besitzt. Dies ist genau dann der Fall, wenn $\det A = ad - bc \neq 0$ ist (Übung). Ist $\ker \varphi = 0$, so ist im φ 2-dimensional, also gleich K^2 , d.h.: φ ist Isomorphismus.

b) Es sei V der Vektorraum der Polynomfunktionen $f : K \rightarrow K$, wobei K ein Körper der Charakteristik Null ist. Dann ist das Ableiten

$$f \mapsto f' =: D(f)$$

eine lineare Abbildung $D : V \rightarrow V$. D ist surjektiv und besitzt ein lineares Rechtsinverses $I : V \rightarrow V$. Mit der Notation in Beispiel 8.12 gilt $D(f_n) = n f_{n-1}$. Damit ist klar, dass durch $I(f_n) := \frac{1}{n+1} f_{n+1}$ ein Rechtsinverses von D definiert wird. Es gilt $\ker D = K f_0$ ist eindimensional. Anders als bei endlichdimensionalen Vektorräumen braucht eine surjektive lineare Abbildung eines unendlichdimensionalen Vektorraums in sich nicht notwendig injektiv zu sein.

Übungen

1. Es sei V ein 2-dimensionaler K -Vektorraum, $\varphi : V \rightarrow V$ sei linear. Zeigen Sie: Es gibt Elemente $a, b \in K$, so dass für alle $x \in V$ gilt:

$$\varphi(\varphi(x)) + a\varphi(x) + bx = 0.$$

Hinweis: Behandeln Sie zunächst den Fall, dass ein $x \in V$ existiert, so dass x und $\varphi(x)$ linear unabhängig sind, und benutzen Sie Lemma 9.3.

2. Es sei $\varphi : \mathbb{R}^4 \rightarrow \mathbb{R}^3$ die lineare Abbildung mit

$$\varphi(e_1) = (1, 1, 0), \quad \varphi(e_2) = (1, 1, 0), \quad \varphi(e_3) = (1, 0, 1), \quad \varphi(e_4) = (2, 1, 1).$$

Bestimmen Sie eine Basis von $\ker \varphi \subset \mathbb{R}^4$ und von $\operatorname{im} \varphi \subset \mathbb{R}^3$.

3. Es sei V ein endlichdimensionaler K -Vektorraum, und W_1, W_2 seien von Null verschiedene komplementäre Untervektorräume von V , also $V = W_1 \oplus W_2$. $\varphi_1 : W_1 \rightarrow V'$, $\varphi_2 : W_2 \rightarrow V'$ seien lineare Abbildungen. Gibt es eine lineare Abbildung $\varphi : V \rightarrow V'$, die sowohl φ_1 als auch φ_2 fortsetzt? Falls ja, ist φ eindeutig bestimmt?
4. Es seien W_1, W_2 endlichdimensionale Untervektorräume eines Vektorraums V . $W_1 \times W_2$ ist dann ein Vektorraum mit den Verknüpfungen

$$\begin{aligned} (x_1, x_2) + (y_1, y_2) &:= (x_1 + y_1, x_2 + y_2) \\ a(x_1, x_2) &:= (ax_1, ax_2). \end{aligned}$$

Zeigen Sie:

$$\varphi : W_1 \times W_2 \rightarrow V, \quad \varphi(x_1, x_2) := x_1 + x_2$$

ist eine lineare Abbildung. Bestimmen Sie Kern und Bild von φ und leiten Sie aus Satz 9.14 erneut die Dimensionsformel 8.14 her.

10 Der Hom-Funktor und Dualräume

Es sei K ein Körper.

Definition 10.1 Es seien V, W K -Vektorräume. Dann sei

$$\text{Hom}(V, W) = \text{Hom}_K(V, W) = \{\varphi : V \rightarrow W \mid \varphi \text{ ist linear} \}$$

die Menge aller Homomorphismen von V nach W .

$\text{Hom}(V, W)$ trägt in natürlicher Weise eine K -Vektorraumstruktur:

a) Für $\varphi, \psi \in \text{Hom}(V, W)$ wird die Summe $\varphi + \psi : V \rightarrow W$ durch

$$(\varphi + \psi)(x) := \varphi(x) + \psi(x) \text{ für alle } x \in V$$

definiert. Es gilt dann $\varphi + \psi \in \text{Hom}(V, W)$, denn

$$\begin{aligned} (\varphi + \psi)(ax + by) &= \varphi(ax + by) + \psi(ax + by) \\ &= a\varphi(x) + b\varphi(y) + a\psi(x) + b\psi(y) \\ &= a(\varphi(x) + \psi(x)) + b(\varphi(y) + \psi(y)) \\ &= a(\varphi + \psi)(x) + b(\varphi + \psi)(y). \end{aligned}$$

Damit haben wir eine Addition

$$+ : \text{Hom}(V, W) \times \text{Hom}(V, W) \rightarrow \text{Hom}(V, W)$$

erklärt.

b) Für $\varphi \in \text{Hom}(V, W)$, $a \in K$ sei $a\varphi : V \rightarrow W$ durch

$$(a\varphi)(x) := a(\varphi(x)) \text{ für alle } x \in V$$

erklärt. Es gilt dann $a\varphi \in \text{Hom}(V, W)$, denn

$$\begin{aligned} (a\varphi)(bx + cy) &= a(\varphi(bx + cy)) = a(b\varphi(x) + c\varphi(y)) \\ &= ab\varphi(x) + ac\varphi(y) = b(a\varphi)(x) + c(a\varphi)(y). \end{aligned}$$

Damit ist eine skalare Multiplikation

$$\cdot : K \times \text{Hom}(V, W) \rightarrow \text{Hom}(V, W)$$

erklärt. Hierfür wurde wesentlich benutzt, dass die Multiplikation von K kommutativ ist. Man sieht leicht, dass für $(\text{Hom}(V, W), +, \cdot)$ die acht Vektorraumaxiome erfüllt sind. Um klar zu machen, dass der Beweis trivial ist, verifizieren wir etwa das Assoziativgesetz der Addition:

$$\begin{aligned} ((\varphi_1 + \varphi_2) + \varphi_3)(x) &= (\varphi_1 + \varphi_2)(x) + \varphi_3(x) \\ &= (\varphi_1(x) + \varphi_2(x)) + \varphi_3(x) = \varphi_1(x) + (\varphi_2(x) + \varphi_3(x)) \\ &= \varphi_1(x) + (\varphi_2 + \varphi_3)(x) = (\varphi_1 + (\varphi_2 + \varphi_3))(x). \end{aligned}$$

Das Nullelement in $\text{Hom}(V, W)$ ist die Nullabbildung $0 : V \rightarrow W$ ($x \mapsto 0$ für alle $x \in V$). Für $\varphi \in \text{Hom}(V, W)$ ist $-\varphi$ durch $(-\varphi)(x) := -\varphi(x)$ definiert.

□

Wir haben also

Satz 10.2 $\text{Hom}(V, W)$ ist ein K -Vektorraum.

Definition 10.3 Es sei V ein K -Vektorraum. Der K -Vektorraum $V^* = \text{Hom}(V, K)$ heißt **Dualraum** von V . Die Elemente $\varphi \in V^*$ heißen **Linearformen** auf V .

Beispiel 10.4 Es sei V der \mathbb{R} -Vektorraum der Polynomfunktionen $f : \mathbb{R} \rightarrow \mathbb{R}$.

- a) Für eine feste Zahl $a \in \mathbb{R}$ ist durch $\varphi_a(f) := f(a)$ für $f \in V$ eine Linearform $\varphi_a : V \rightarrow \mathbb{R}$ definiert, denn für $f, g \in V$ ist

$$\varphi_a(f + g) = (f + g)(a) = f(a) + g(a) = \varphi_a(f) + \varphi_a(g),$$

und für $b \in \mathbb{R}$ und $f \in V$ ist

$$\varphi_a(bf) = (bf)(a) = bf(a) = b\varphi_a(f).$$

- b) Es seien $a, b \in \mathbb{R}$, $a < b$. Dann ist auch

$$f \mapsto \varphi(f) := \int_a^b f(x) dx$$

eine Linearform $\varphi \in V^*$. Dies kann man hier leicht beweisen, wenn man das Integral rein algebraisch einführt durch

$$\int_a^b \left(\sum_{\nu=0}^n a_\nu x^\nu \right) dx = \sum_{\nu=0}^n a_\nu \frac{b^{\nu+1} - a^{\nu+1}}{\nu+1}.$$

- c) $f \mapsto \psi(f) := f(0)^2$ ist eine Abbildung $\psi : V \rightarrow \mathbb{R}$, aber natürlich nicht linear!
 d) Es sei $D : V \rightarrow V$ der Ableitungsoperator. Dann ist $\psi_a := \varphi_a \circ D \in V^*$, denn da D und φ_a lineare Abbildungen sind, ist auch $\varphi_a \circ D$ linear. Als kleine Übung mit dem neuen Begriff der Linearform wollen wir zeigen: φ_a, ψ_a sind linear unabhängig in V^* .

Beweis. Es seien $a_1, a_2 \in \mathbb{R}$, und es gelte

$$a_1\varphi_a + a_2\psi_a = 0,$$

d.h. für alle $f \in V$ ist $a_1\varphi_a(f) + a_2\psi_a(f) = 0$, also

$$a_1f(a) + a_2f'(a) = 0 \quad \forall f \in V.$$

Setzt man die beiden Funktionen $f(x) = 1$ und $g(x) = x$ ein, so erhält man

$$\begin{aligned} a_1 \cdot 1 + a_2 \cdot 0 &= 0, \\ a_1 a + a_2 \cdot 1 &= 0, \end{aligned}$$

also $a_1 = a_2 = 0$. □

Definition 10.5 Es seien V und W K -Vektorräume und $\alpha : V \rightarrow W$ sei eine lineare Abbildung. Ist $\varphi \in W^*$, also eine Linearform $\varphi : W \rightarrow K$, so ist auch $\varphi \circ \alpha : V \rightarrow K$ eine Linearform. Man setzt

$$\alpha^*(\varphi) := \varphi \circ \alpha.$$

Dadurch ist eine Abbildung

$$\alpha^* : W^* \rightarrow V^*$$

definiert. α^* heißt die **duale Abbildung** von α .

Man sieht sofort

Lemma 10.6 Ist $\alpha : V \rightarrow W$ linear, so ist auch $\alpha^* : W^* \rightarrow V^*$ linear.

Beweis:

a) Seien $\varphi_1, \varphi_2 \in W^*$. Dann gilt für alle $x \in V$:

$$\begin{aligned} ((\varphi_1 + \varphi_2) \circ \alpha)(x) &= (\varphi_1 + \varphi_2)(\alpha(x)) = \varphi_1(\alpha(x)) + \varphi_2(\alpha(x)) \\ &= (\varphi_1 \circ \alpha)(x) + (\varphi_2 \circ \alpha)(x) = (\varphi_1 \circ \alpha + \varphi_2 \circ \alpha)(x), \end{aligned}$$

also

$$\alpha^*(\varphi_1 + \varphi_2) = (\varphi_1 + \varphi_2) \circ \alpha = \varphi_1 \circ \alpha + \varphi_2 \circ \alpha = \alpha^*(\varphi_1) + \alpha^*(\varphi_2).$$

b) Für $\varphi \in W^*$, $a \in K$ und $x \in V$ gilt:

$$((a\varphi) \circ \alpha)(x) = (a\varphi)(\alpha(x)) = a\varphi(\alpha(x)) = a(\varphi \circ \alpha)(x);$$

und somit

$$\alpha^*(a\varphi) = (a\varphi) \circ \alpha = a(\varphi \circ \alpha) = a\alpha^*(\varphi).$$

□

Satz 10.7 (funktorielle Eigenschaften): Es gilt:

a) Ist V ein K -Vektorraum, so ist

$$(id_V)^* = id_{V^*}.$$

b) Sind V, V', V'' K -Vektorräume und $\alpha : V \rightarrow V'$, $\beta : V' \rightarrow V''$ linear, so gilt

$$(\beta \circ \alpha)^* = \alpha^* \circ \beta^*.$$

Beweis: a) ist klar.

Zu b): Für $\varphi \in V''^*$ gilt:

$$\begin{aligned} (\alpha^* \circ \beta^*)(\varphi) &= \alpha^*(\beta^*(\varphi)) = \alpha^*(\varphi \circ \beta) = (\varphi \circ \beta) \circ \alpha \\ &= \varphi \circ (\beta \circ \alpha) = (\beta \circ \alpha)^*(\varphi), \end{aligned}$$

also ist $(\beta \circ \alpha)^* = \alpha^* \circ \beta^*$.

□

Jetzt betrachten wir endlichdimensionale Vektorräume.

Satz 10.8 Es sei V ein n -dimensionaler K -Vektorraum, und (x_1, \dots, x_n) sei eine Basis von V . $\varphi_i : V \rightarrow \mathbb{R}$ sei die Linearform mit

$$\varphi_i(x_j) = \delta_{ij} := \begin{cases} 1 & \text{falls } i = j \\ 0 & \text{falls } i \neq j \end{cases}$$

für $j = 1, \dots, n$. Dann ist $(\varphi_1, \dots, \varphi_n)$ eine Basis des Dualraumes V^* . Insbesondere gilt also auch: $\dim V = \dim V^*$.

Beweis: Man beachte, dass die Linearformen $\varphi_i \in V^*$ nach Satz 9.4 durch die Bedingung

$$\varphi_i(x_j) = \delta_{ij}$$

eindeutig bestimmt sind.

Wir zeigen zunächst, dass $\varphi_1, \dots, \varphi_n$ linear unabhängig sind. Seien also $a_1, \dots, a_n \in K$, und es sei

$$a_1\varphi_1 + \dots + a_n\varphi_n = 0.$$

Dann gilt also für alle $x \in V$:

$$\sum_{i=1}^n a_i\varphi_i(x) = 0;$$

insbesondere für $x = x_j$ ergibt sich

$$\sum_{i=1}^n a_i\delta_{ij} = 0, \quad \text{d.h. } a_j = 0.$$

Jetzt beweisen wir, dass V^* von $\varphi_1, \dots, \varphi_n$ erzeugt wird. Es sei dazu $\varphi \in V^*$ beliebig. Dann ist $a_i := \varphi(x_i) \in K$ für $i = 1, \dots, n$.

$\psi := a_1\varphi_1 + \dots + a_n\varphi_n$ ist ebenfalls eine Linearform auf V . Für die beiden linearen Abbildungen

$$\varphi, \psi : V \rightarrow K$$

gilt nun:

$$\psi(x_j) = \sum_{i=1}^n a_i\varphi_i(x_j) = \sum_{i=1}^n a_i\delta_{ij} = a_j = \varphi(x_j)$$

für $j = 1, \dots, n$; und somit folgt aus Lemma 9.3: $\varphi = \psi$. □

Definition 10.9

- a) Die Basis $(\varphi_1, \dots, \varphi_n)$ von V^* heißt die zu (x_1, \dots, x_n) **duale Basis**.
- b) δ_{ij} heißt das **Kroneckersymbol**.

Bemerkung 10.10 Besonders in der Physik ist es üblich, statt $\varphi(x)$ die Schreibweise $\langle \varphi | x \rangle$ zu benutzen. Die Abbildung (Auswertungsabbildung)

$$\begin{aligned} V^* \times V &\longrightarrow K \\ (\varphi, x) &\longmapsto \varphi(x) = \langle \varphi | x \rangle \end{aligned}$$

heißt auch **duale Paarung**.

Ist (x_1, \dots, x_n) eine Basis von V und $\varphi \in V^*$, so heißt $\langle \varphi | x_i \rangle = \varphi(x_i)$ der i -te **Fourierkoeffizient** von φ bezüglich der Basis (x_1, \dots, x_n) . Ist $(\varphi_1, \dots, \varphi_n)$ die duale Basis von (x_1, \dots, x_n) , so ergibt sich aus dem Beweis von Satz 10.8:

$$\varphi = \langle \varphi | x_1 \rangle \varphi_1 + \dots + \langle \varphi | x_n \rangle \varphi_n,$$

und für $x = a_1 x_1 + \dots + a_n x_n$ ergibt sich wegen $\langle \varphi_i | x_j \rangle = \delta_{ij}$:

$$\langle \varphi | x \rangle = \langle \varphi | x_1 \rangle a_1 + \dots + \langle \varphi | x_n \rangle a_n.$$

Beispiel 10.11 Es sei $\alpha : K^2 \rightarrow K^2$ linear, $e_1 = (1, 0)$, $e_2 = (0, 1)$, $\alpha(e_1) = (a, c) = a e_1 + c e_2$, $\alpha(e_2) = (b, d) = b e_1 + d e_2$. Es sei weiter (φ_1, φ_2) die duale Basis von (e_1, e_2) , also $\varphi_i(e_j) = \delta_{ij}$. Dann ist

$$\begin{aligned} \alpha^*(\varphi_1) &= \varphi_1 \circ \alpha = \langle \varphi_1 \circ \alpha | e_1 \rangle \varphi_1 + \langle \varphi_1 \circ \alpha | e_2 \rangle \varphi_2 \\ &= \langle \varphi_1 | \alpha(e_1) \rangle \varphi_1 + \langle \varphi_1 | \alpha(e_2) \rangle \varphi_2 \\ &= \langle \varphi_1 | a e_1 + c e_2 \rangle \varphi_1 + \langle \varphi_1 | b e_1 + d e_2 \rangle \varphi_2 \\ &= a \varphi_1 + b \varphi_2; \end{aligned}$$

und analog

$$\alpha^*(\varphi_2) = c \varphi_1 + d \varphi_2.$$

α wird bezüglich (e_1, e_2) durch die Matrix $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ gegeben.

α^* wird bezüglich (φ_1, φ_2) durch die Matrix $A^t = \begin{pmatrix} a & c \\ b & d \end{pmatrix}$ gegeben.

Definition 10.12 Es seien V_1, \dots, V_{n+1} K -Vektorräume und $\alpha_i : V_i \rightarrow V_{i+1}$ lineare Abbildungen für $i = 1, \dots, n$. Dann nennt man

$$V_1 \xrightarrow{\alpha_1} V_2 \xrightarrow{\alpha_2} V_3 \longrightarrow \dots \longrightarrow V_n \xrightarrow{\alpha_n} V_{n+1}$$

eine **Sequenz** linearer Abbildungen.

Eine solche Sequenz heißt **exakt**, wenn

$$\ker \alpha_{i+1} = \operatorname{im} \alpha_i$$

für $i = 1, \dots, n - 1$ gilt.

Beispiel 10.13 Es sei V ein K -Vektorraum und $U, W \subset V$ seien Untervektorräume, so dass $U \cap W = 0$ und $V = U + W$. Dann hat jeder Vektor $x \in V$ eine eindeutige Darstellung

$$x = x_1 + x_2 \text{ mit } x_1 \in U \text{ und } x_2 \in W.$$

Wir setzen $\pi(x) := x_2$. Dadurch ist eine lineare Abbildung $\pi : V \rightarrow W$ definiert. Es sei weiter $\iota : U \rightarrow V$ die Inklusion. Dann ist

$$0 \longrightarrow U \xrightarrow{\iota} V \xrightarrow{\pi} W \longrightarrow 0$$

eine exakte Sequenz. Dabei bezeichnet 0 den Nullvektorraum. $0 \rightarrow U, W \rightarrow 0$ sind Nullabbildungen.

Zum Beweis der Exaktheit:

1. $\ker \iota = 0 = \text{im}(0 \rightarrow U)$
2. Da $\pi : V \rightarrow W$ surjektiv ist, denn für $w \in W$ ist $\pi(w) = w$, gilt $\ker(W \rightarrow 0) = W = \text{im } \pi$.
3. Ist $x = x_1 + x_2$, $x_1 \in U$, $x_2 \in W$, so gilt

$$\pi(x) = 0 \iff x_2 = 0 \iff x = x_1 = \iota(x_1),$$

d.h.: $\ker \pi = \text{im } \iota$.

Lemma 10.14 Es sei V ein endlichdimensionaler Vektorraum, $U \subset V$ ein Untervektorraum, $i : U \rightarrow V$ die Inklusionsabbildung. Dann ist $i^* : V^* \rightarrow U^*$ die Abbildung, die jeder Linearform $\varphi : V \rightarrow K$ ihre Einschränkung $\varphi|_U = \varphi \circ i : U \rightarrow K$ zuordnet. Es gilt: i^* ist surjektiv.

Beweis: Wir müssen zeigen, dass jede Linearform $\psi : U \rightarrow K$ eine lineare Fortsetzung $\varphi : V \rightarrow K$ besitzt. Dazu wähle man eine Basis (x_1, \dots, x_m) von U und ergänze diese zu einer Basis (x_1, \dots, x_n) von V . Jetzt sei $\varphi : V \rightarrow K$ die lineare Abbildung mit

$$\varphi(x_i) = \begin{cases} \psi(x_i) & \text{für } i = 1, \dots, m \\ 0 & \text{für } i = m+1, \dots, n. \end{cases}$$

Ersichtlich gilt: $\varphi \circ i = \psi$. □

Jetzt beweisen wir folgenden wichtigen

Satz 10.15 Es seien V_1, V_2, V_3 endlichdimensionale K -Vektorräume, und es sei

$$V_1 \xrightarrow{\alpha} V_2 \xrightarrow{\beta} V_3$$

eine exakte Sequenz. Dann ist auch

$$V_3^* \xrightarrow{\beta^*} V_2^* \xrightarrow{\alpha^*} V_1^*$$

exakt.

Beweis:

1. *Behauptung:*

$$\text{im } \beta^* \subset \ker \alpha^*.$$

Beweis: Da $\text{im } \alpha \subset \ker \beta$ gilt, ist $\beta \circ \alpha = 0$, und somit erhält man eine Faktorisierung

$$\begin{array}{ccc} V_1 & \xrightarrow{\beta \circ \alpha} & V_3 \\ & \searrow \gamma & \nearrow i \\ & 0 & \end{array}$$

$$(\beta \circ \alpha = i \circ \gamma).$$

Es folgt

$$\alpha^* \circ \beta^* = (\beta \circ \alpha)^* = (i \circ \gamma)^* = \gamma^* \circ i^* = 0,$$

denn i^* ist die Nullabbildung. Also gilt $\alpha^* \circ \beta^* = 0$, d.h. $\text{im } \beta^* \subset \ker \alpha^*$.

2. *Behauptung:*

$$\ker \alpha^* \subset \text{im } \beta^*.$$

Beweis: Es sei $\varphi \in \ker \alpha^*$, also $\varphi : V_2 \rightarrow K$ linear mit $\varphi \circ \alpha = 0$. Folglich ist

$$U := \ker \beta = \text{im } \alpha \subset \ker \varphi.$$

Es sei nun $W \subset V_2$ ein zu U komplementärer Untervektorraum von V_2 , also

$$U \cap W = 0, \quad V_2 = U + W.$$

Es sei $\pi : V_2 \rightarrow W$ die Projektion $\pi(x) = x_2$, wenn $x = x_1 + x_2$, $x_1 \in U$, $x_2 \in W$. Es gilt dann

$$\varphi(x) = \varphi(x_1 + x_2) = \varphi(x_1) + \varphi(x_2) = \varphi(x_2) = \varphi(\pi(x)),$$

denn $x_1 \in U \subset \ker \varphi$.

Setzt man $\bar{\varphi} := \varphi|_W$, so gilt also

$$\varphi = \bar{\varphi} \circ \pi.$$

Ebenso gilt für $\bar{\beta} = \beta|_W$

$$\beta = \bar{\beta} \circ \pi.$$

$\bar{\beta}$ ist injektiv, denn

$$\ker(\bar{\beta}) = W \cap \ker \beta = W \cap U = 0$$

nach Wahl von W .

Bezeichnet man mit $\iota : \text{im } \beta \rightarrow V_3$ die Inklusion, so erhält man einen Isomorphismus

$$\tilde{\beta} : W \rightarrow \text{im } \beta$$

mit

$$\iota \circ \tilde{\beta} = \bar{\beta}.$$

Es sei $\psi : V_3 \rightarrow K$ eine lineare Fortsetzung von $\bar{\varphi} \circ \tilde{\beta}^{-1} : \text{im } \beta \rightarrow K$. Es gilt dann

$$\psi \circ \iota = \bar{\varphi} \circ \tilde{\beta}^{-1}, \quad \text{also } \psi \circ \iota \circ \tilde{\beta} = \bar{\varphi}$$

und somit

$$\varphi = \bar{\varphi} \circ \pi = \psi \circ \iota \circ \tilde{\beta} \circ \pi = \psi \circ \bar{\beta} \circ \pi = \psi \circ \beta,$$

d.h. $\varphi = \beta^*(\psi)$.

Somit ist $\varphi \in \text{im } \beta^*$. □

Als wichtige Folgerung erhalten wir

Korollar 10.16 Es seien V, W endlichdimensionale K -Vektorräume. $\alpha : V \rightarrow W$ sei linear. $\alpha^* : W^* \rightarrow V^*$ sei die duale Abbildung. Dann gilt:

- (i) α ist injektiv $\Rightarrow \alpha^*$ ist surjektiv.

(ii) α ist surjektiv $\Rightarrow \alpha^*$ ist injektiv.

Beweis: zu (i): α ist injektiv $\Rightarrow 0 \rightarrow V \xrightarrow{\alpha} W$ ist exakt $\Rightarrow W^* \xrightarrow{\alpha^*} V^* \rightarrow 0$ ist exakt $\Rightarrow \alpha^*$ ist surjektiv.

zu (ii): α ist surjektiv $\Rightarrow V \xrightarrow{\alpha} W \rightarrow 0$ ist exakt $\Rightarrow 0 \rightarrow W^* \xrightarrow{\alpha^*} V^*$ ist exakt $\Rightarrow \alpha^*$ ist injektiv. \square

Korollar 10.17 Es seien V, W endlichdimensionale K -Vektorräume. $\alpha : V \rightarrow W$ sei lineare Abbildung. Dann gilt:

$$\operatorname{rg} \alpha = \operatorname{rg} \alpha^*.$$

Beweis: Wir betrachten die Faktorisierung

$$\begin{array}{ccc} V & \xrightarrow{\alpha} & W \\ & \searrow \alpha' & \nearrow i \\ & \operatorname{im} \alpha & \end{array}$$

von α . Dann ist

$$\begin{array}{ccc} W^* & \xrightarrow{\alpha^*} & V^* \\ & \searrow i^* & \nearrow \alpha'^* \\ & (\operatorname{im} \alpha)^* & \end{array}$$

eine Faktorisierung von α^* . Nach 10.16 (i) ist i^* surjektiv, und nach 10.16 (ii) ist α'^* injektiv, also ist

$$\operatorname{rg} \alpha^* = \operatorname{rg} i^* = \dim(\operatorname{im} \alpha)^* = \dim(\operatorname{im} \alpha) = \operatorname{rg} \alpha.$$

\square

Definition 10.18 Ist V ein K -Vektorraum, so heißt

$$V^{**} := (V^*)^*$$

der **Bidualraum** von V . Für $x \in V$ wird $\hat{x} \in V^{**}$ definiert durch

$$\hat{x}(\varphi) := \varphi(x) \quad \forall \varphi \in V^*.$$

In der Tat gilt:

$$\hat{x}(\varphi_1 + \varphi_2) = (\varphi_1 + \varphi_2)(x) = \varphi_1(x) + \varphi_2(x) = \hat{x}(\varphi_1) + \hat{x}(\varphi_2)$$

und

$$\hat{x}(a\varphi) = (a\varphi)(x) = a\varphi(x) = a\hat{x}(\varphi),$$

also ist $\hat{x} : V^* \rightarrow K$ linear, d.h. $\hat{x} \in V^{**}$.

Mit $\tau_V : V \rightarrow V^{**}$ bezeichnen wir die kanonische Abbildung $x \mapsto \hat{x}$, also $\tau_V(x) = \hat{x}$.

Satz 10.19 a) Ist V ein endlichdimensionaler K -Vektorraum, so ist $\tau_V : V \rightarrow V^{**}$ ein Isomorphismus von Vektorräumen.

b) Ist $\alpha : V \rightarrow W$ eine lineare Abbildung von K -Vektorräumen, so gilt:

$$\tau_W \circ \alpha = \alpha^{**} \circ \tau_V.$$

Beweis: Zu a):

1. Behauptung: $\tau_V : V \rightarrow V^{**}$ ist linear.

Das ist Spielerei. Seien $x, y \in V$, $a, b \in K$ und $\varphi \in V^*$. Dann gilt:

$$\begin{aligned} \tau_V(ax + by)(\varphi) &= \varphi(ax + by) = a\varphi(x) + b\varphi(y) \\ &= a\tau_V(x)(\varphi) + b\tau_V(y)(\varphi) = (a\tau_V(x) + b\tau_V(y))(\varphi), \end{aligned}$$

also

$$\tau_V(ax + by) = a\tau_V(x) + b\tau_V(y).$$

□

2. Behauptung: τ_V ist injektiv.

Dazu zeigen wir, dass $\tau_V(x) \neq 0$ ist, wenn $x \neq 0$ ist. Sei also $x \in V \setminus \{0\}$. Dann ist x linear unabhängig, und wir finden eine Basis x_1, \dots, x_n von V mit $x_1 = x$. Nun sei z.B. $\varphi \in V^*$ die Linearform mit $\varphi(x_i) = \delta_{1i}$. Es gilt dann:

$$\tau_V(x)(\varphi) = \varphi(x) = \delta_{11} = 1 \neq 0,$$

also auch $\tau_V(x) \neq 0$.

□

3. Behauptung: τ_V ist bijektiv.

Da $\ker \tau_V = 0$, gilt nach Satz 10.16 $\dim(\operatorname{im} \tau_V) = \dim V$ und da nach Satz 10.8 $\dim V = \dim V^*$ gilt, muss im $\tau_V = V^{**}$ gelten. τ_V ist also surjektiv.

□

zu b): Für $x \in V$, $\psi \in W^*$ gilt:

$$\begin{aligned} ((\tau_W \circ \alpha)(x))(\psi) &= (\tau_W(\alpha(x)))(\psi) = \psi(\alpha(x)) = (\psi \circ \alpha)(x) \\ ((\alpha^{**} \circ \tau_V)(x))(\psi) &= (\alpha^{**}(\tau_V(x)))(\psi) = (\tau_V(x) \circ \alpha^*)(\psi) \\ &= \tau_V(x)(\alpha^*(\psi)) = \tau_V(x)(\psi \circ \alpha) = (\psi \circ \alpha)(x). \end{aligned}$$

□

$\tau_V : V \rightarrow V^{**}$ heißt der **natürliche Isomorphismus** von V nach V^{**} . Jeden Vektor $x \in V$ kann man als Linearform auf dem Dualraum V^* auffassen:

$$(\varphi \mapsto \varphi(x)).$$

Jede Linearform $\lambda : V^* \rightarrow K$ ist von dieser Form: Zu λ gibt es genau einen Vektor $x \in V$, so dass $\lambda(\varphi) = \varphi(x)$ für alle $\varphi \in V^*$.

Übungen

- (a) Es sei V der Vektorraum der Polynomfunktionen $f : \mathbb{R} \rightarrow \mathbb{R}$. Für $f \in V$ sei f' die Ableitung von f und

$$\varphi_1(f) := 2f(0) + f'(1) \quad \text{und} \quad \varphi_2(f) := \int_0^1 f(x) dx.$$

Zeigen Sie, dass $\varphi_1, \varphi_2 \in V^*$ und dass φ_1 und φ_2 linear unabhängig sind.

(b) Finden Sie eine lineare Abbildung $\alpha : \mathbb{R}^2 \rightarrow V$, so dass $\alpha^*(\varphi_1), \alpha^*(\varphi_2)$ linear unabhängig in $(\mathbb{R}^2)^*$ sind.

(c) Finden Sie ein Erzeugendensystem des Untervektorraums

$$W = \ker \varphi_1 \cap \ker \varphi_2 \cap V_n \subset V,$$

wobei V_n der Untervektorraum der Polynomfunktionen vom Grad $\leq n$ ist.

2. Es sei V ein n -dimensionaler \mathbb{R} -Vektorraum, und (x_1, \dots, x_n) sei eine Basis von V . $\varphi, \psi \in V^*$ seien die Linearformen mit

$$\varphi(x_\nu) = \nu \quad \text{und} \quad \psi(x_\nu) = (-1)^\nu \nu \quad \text{für} \quad \nu = 1, \dots, n.$$

Stellen Sie $\varphi, \psi, \varphi + \psi, \varphi - \psi$ als Linearkombination in der zu (x_1, \dots, x_n) dualen Basis dar.

Bestimmen Sie alle linearen Abbildungen $\alpha : V \rightarrow V$ mit $\alpha^*(\varphi) = \psi$.

3. Es seien V, W endlichdimensionale K -Vektorräume. Dann ist $\alpha \mapsto \alpha^*$ ein K -Vektorraumisomorphismus

$$\text{Hom}(V, W) \longrightarrow \text{Hom}(W^*, V^*).$$

4. Mit Hilfe von Satz 10.15 zeige man:

Es sei V ein endlichdimensionaler K -Vektorraum. $U, \tilde{U} \subset V$ seien komplementäre Unterräume, also $V = U \oplus \tilde{U}$. $\pi : V \rightarrow \tilde{U}$ sei die Projektion $\pi(u + \tilde{u}) = \tilde{u}$, $u \in U, \tilde{u} \in \tilde{U}$. Zeigen Sie, dass $\text{im}(\pi^*) = U^\perp$ ist, wobei

$$U^\perp := \{\varphi \in V^* \mid \varphi(x) = 0 \text{ für alle } x \in U\}.$$

Wie ist der Zusammenhang zwischen $\dim U$ und $\dim U^\perp$?

Ist $W \subset V^*$ ein Untervektorraum, so ist

$$W^\perp := \{x \in V \mid \varphi(x) = 0 \text{ für alle } \varphi \in W\}$$

ein Untervektorraum von V . Es gilt: $(W^\perp)^\perp = W$.

5. Es seien P, Q, R, S vier paarweise verschiedene kollineare Punkte in \mathbb{R}^2 (Punkte heißen kollinear, wenn sie auf einer Geraden liegen.)

Dann sind die Vektoren

$$\overrightarrow{PQ} := Q - P, \quad \overrightarrow{RS} := S - R$$

linear abhängig. Es gilt also

$$\overrightarrow{PQ} = a \overrightarrow{RS} \text{ mit } a \in \mathbb{R}.$$

Die Zahl a heißt das Verhältnis von \overrightarrow{PQ} zu \overrightarrow{RS} und wird mit

$$\frac{PQ}{RS}$$

bezeichnet.

Ist nun $\varphi : \mathbb{R}^2 \rightarrow \mathbb{R}$ eine Linearform mit $\varphi(R) \neq \varphi(S)$, so gilt

$$\frac{PQ}{RS} = \frac{\varphi(Q) - \varphi(R)}{\varphi(S) - \varphi(R)}.$$

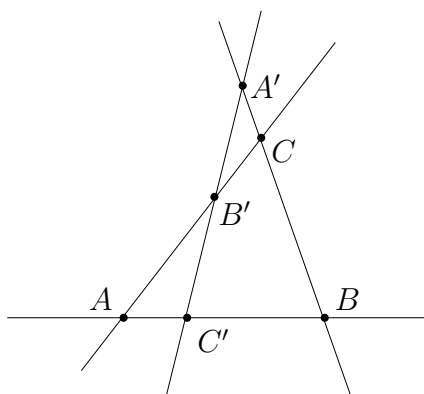
6. Es seien A, B, C drei nicht kollineare Punkte in \mathbb{R}^2 . Weiter seien

$$A' \in g_{B,C}, \quad B' \in g_{A,C}, \quad C' \in g_{A,B}$$

paarweise verschieden, und es sei $\{A, B, C\} \cap \{A', B', C'\} = \emptyset$. Dabei bezeichnet $g_{P,Q} = \{P + t(Q - P) \mid t \in \mathbb{R}\}$ die Verbindungsgerade von P und Q in \mathbb{R}^2 .

Zeigen Sie: A', B', C' sind genau dann kollinear, wenn gilt:

$$\frac{A'B}{A'C} \cdot \frac{B'C}{B'A} \cdot \frac{C'A}{C'B} = 1.$$

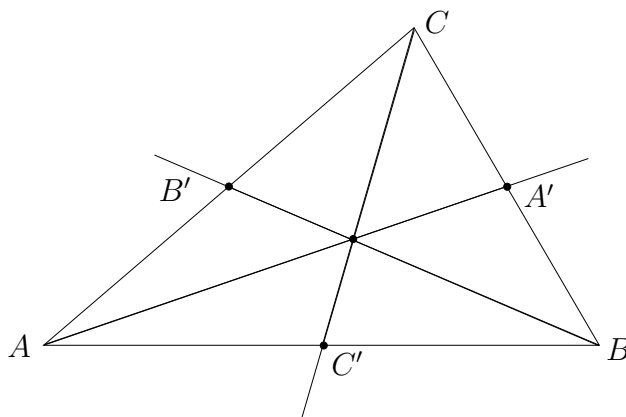


Hinweis: Verwenden Sie eine geeignete Linearform $\varphi : \mathbb{R}^2 \rightarrow \mathbb{R}$ und Aufgabe 5.

7. Die Voraussetzungen seien wie in Aufgabe 6.

Die Geraden $g_{A,A'}$, $g_{B,B'}$, $g_{C,C'}$ schneiden sich in einem Punkt, genau dann, wenn gilt:

$$\frac{A'B}{A'C} \cdot \frac{B'C}{B'A} \cdot \frac{C'A}{C'B} = -1.$$



Hinweis: Wie bei Aufgabe 6.

11 Lineare Abbildungen und Matrizen

Definition 11.1 Es seien $m, n \in \mathbb{N}$, $m, n > 0$. X sei eine Menge. Eine $m \times n$ -**Matrix** über X (oder: eine $m \times n$ -Matrix mit Koeffizienten in X) ist eine Tabelle

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

von Elementen $a_{ij} \in X$.

Man beachte, dass ij ein Doppelindex ist. Hier bedeutet also 11 in a_{11} nicht "elf", sondern "eins, eins". Da immer klar sein wird, was gemeint ist, werden wir stets a_{ij} , statt etwa $a_{i,j}$ schreiben!

$$a_i := (a_{i1}, \dots, a_{in})$$

heißt die i -te **Zeile** von A . a_i ist eine $1 \times n$ -Matrix über X .

$$a^j := \begin{pmatrix} a_{1j} \\ \vdots \\ a_{mj} \end{pmatrix} = [a_{1j}, a_{2j}, \dots, a_{mj}]$$

heißt die j -te **Spalte** von A . a^j ist eine $m \times 1$ -Matrix über X . Wir schreiben dann

$$A = (a^1, \dots, a^n) = \begin{pmatrix} a_1 \\ \vdots \\ a_m \end{pmatrix}.$$

Mit $X^{m,n}$ bezeichnen wir die Menge aller $m \times n$ -Matrizen über X .

Beispiel 11.2 Für $X = \{0, 1\}$ ist

$$X^{2,2} = \left\{ \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 1 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 1 \end{pmatrix}, \dots \right\}$$

eine Menge mit 16 Elementen. Ist X eine Menge mit q Elementen, so hat $X^{m,n}$ q^{mn} Elemente.

Es sei K ein Körper und $m, n \in \mathbb{N}$, $m, n \geq 1$. Wir wollen zunächst die linearen Abbildungen

$$\varphi : K^n \rightarrow K^m$$

studieren. Für $x = (x_1, \dots, x_n) \in K^n$ sei $y := \varphi(x)$; also:

$$y = (y_1, \dots, y_m) \in K^m$$

und

$$(y_1, \dots, y_m) = \varphi(x_1, \dots, x_n).$$

Wie kann man die Komponenten y_i von y aus den Komponenten x_j von x berechnen? Es sei (e_1, \dots, e_n) die Standardbasis von K^n und (e'_1, \dots, e'_m) die Standardbasis von K^m . Dann ist

$$y = \sum_{i=1}^m y_i e'_i, \quad x = \sum_{j=1}^n x_j e_j$$

und somit

$$\varphi(x) = \sum_{j=1}^n x_j \varphi(e_j),$$

denn φ ist ja linear. Es gilt also

$$\sum_{j=1}^n x_j \varphi(e_j) = \sum_{i=1}^m y_i e'_i. \quad (35)$$

Es seien nun $a_{ij} \in K$ die Elemente mit

$$\varphi(e_j) = (a_{1j}, \dots, a_{mj}) = \sum_{i=1}^m a_{ij} e'_i \quad \text{für } j = 1, \dots, n. \quad (36)$$

Dann erhält man durch Einsetzen von (36) in (35):

$$\sum_{j=1}^n \sum_{i=1}^m x_j a_{ij} e'_i = \sum_{i=1}^m y_i e'_i.$$

Ordnet man die Summanden auf der linken Seite um, so erhält man

$$\sum_{i=1}^m \left(\sum_{j=1}^n a_{ij} x_j \right) e'_i = \sum_{i=1}^m y_i e'_i;$$

und hieraus ergibt sich durch Koeffizientenvergleich (e'_1, \dots, e'_m sind ja linear unabhängig!)

$$y_i = \sum_{j=1}^n a_{ij} x_j \quad \text{für } i = 1, \dots, m. \quad (37)$$

Durch die Formeln (37) ist also die Abbildung $x \mapsto y = \varphi(x)$ mit Hilfe der $m \times n$ -Matrix $A = (a_{ij}) \in K^{m,n}$ beschrieben. Es erweist sich nun als vorteilhaft im Hinblick auf den Matrizenkalkül die Vektoren in K^m und K^n als Spaltenvektoren, d. h. als $n \times 1$ -Matrizen, aufzufassen, was wir ab jetzt tun wollen.

Die Gleichungen (37) können wir dann in der Form

$$\begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix} = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

schreiben. Dabei definieren wir das Produkt einer Matrix mit einem Spaltenvektor wie folgt:

Definition 11.3 Für eine $m \times n$ -Matrix $A = (a_{ij}) \in K^{m,n}$ und eine $n \times 1$ -Matrix $b = (b_j) \in K^{n,1}$ (d.h. für einen Spaltenvektor b) wird das Produkt

$$A b$$

definiert als die $m \times 1$ -Matrix $c = (c_i) \in K^{m,1}$ mit

$$c_i := \sum_{j=1}^n a_{ij} b_j \quad \text{für } i = 1, \dots, m.$$

Ab jetzt wollen wir die Elemente aus K^n als Spaltenvektoren schreiben. Mit e^i bezeichnen wir den i -ten Einheitsspaltenvektor in K^n .

Satz 11.4 Jede lineare Abbildung $\varphi : K^n \rightarrow K^m$ ist von der Form $x \mapsto Ax$, wobei $x \in K^n$ als Spaltenvektor $x = [x_1, \dots, x_n]$ geschrieben wird und $A = (a^1, \dots, a^n)$ die Matrix mit den Spalten $a^i = \varphi(e^i)$ ist, $e^i := [0, \dots, 1, \dots, 0]$. Hierdurch wird eine bijektive Abbildung

$$\begin{aligned} \Phi : \text{Hom}(K^n, K^m) &\longrightarrow K^{m,n} \\ \varphi &\longmapsto A = (\varphi(e^1), \dots, \varphi(e^n)) \end{aligned}$$

mit Umkehrabbildung

$$\begin{aligned} \Psi : K^{m,n} &\longrightarrow \text{Hom}(K^n, K^m) \\ A &\longmapsto (x \mapsto Ax) \end{aligned}$$

definiert.

Beweis:

- a) Sei $A \in K^{m,n}$. Dann ist $\varphi : K^n \rightarrow K^m$ mit $\varphi(x) = Ax$ für $x \in K^n$ eine lineare Abbildung, und es gilt:

$$\varphi(e^i) = A \begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix} = a^i$$

ist die i -te Spalte von A . Also ist A auch die durch φ definierte $m \times n$ -Matrix, d.h. $\Phi \circ \Psi = id$.

- b) Ist $\varphi \in \text{Hom}(K^n, K^m)$ und A die Matrix mit i -ter Spalte $a^i = \varphi(e^i)$, so gilt $\varphi(x) = Ax$ für alle $x \in K^n$, also $\Psi \circ \Phi = id$. \square

Wir dürfen $\text{Hom}(K^n, K^m)$ und die Menge $K^{m,n}$ der $m \times n$ -Matrizen identifizieren:

$$\text{Hom}(K^n, K^m) = K^{m,n}.$$

Beispiel 11.5 a) $\varphi : \mathbb{R}^3 \rightarrow \mathbb{R}^4$ mit

$$\varphi \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x + 2z \\ y - x \\ z + y \\ z + x \end{pmatrix}$$

definiert die 4×3 -Matrix

$$A = \begin{pmatrix} 1 & 0 & 2 \\ -1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \end{pmatrix},$$

und es gilt dann:

$$\varphi \begin{pmatrix} x \\ y \\ z \end{pmatrix} = A \begin{pmatrix} x \\ y \\ z \end{pmatrix},$$

kurz:

$$\varphi = A.$$

b) $A = (1 \ 2 \ 3)$ definiert die lineare Abbildung $\varphi : \mathbb{R}^3 \rightarrow \mathbb{R}^1$ mit

$$\varphi \begin{pmatrix} x \\ y \\ z \end{pmatrix} = A \begin{pmatrix} x \\ y \\ z \end{pmatrix} = (1 \ 2 \ 3) \begin{pmatrix} x \\ y \\ z \end{pmatrix} = x + 2y + 3z.$$

c) $A = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$ definiert die lineare Abbildung

$$\varphi : \mathbb{R}^1 \rightarrow \mathbb{R}^3, \quad \varphi(x) = A(x) = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} x = \begin{pmatrix} x \\ 2x \\ 3x \end{pmatrix}.$$

Aus b) und c) entnehmen wir:

$$K^n = \left\{ x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \mid x_i \in K \right\} = K^{n,1}$$

und

$$(K^n)^* = \text{Hom}(K^n, K) = K^{1,n} = \{ a = (a_1, \dots, a_n) \mid a_i \in K \}.$$

Die duale Paarung

$$(K^n)^* \times K^n \rightarrow K$$

ist die Matrizenmultiplikation

$$(a, x) \mapsto a \cdot x = a_1 x_1 + \dots + a_n x_n.$$

Definition 11.6 Die Elemente von K^n schreiben wir als Spalten und die Elemente des Dualraums $(K^n)^*$ als Zeilen. Die Abbildung

$$x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \mapsto x^t = (x_1, \dots, x_n)$$

(lies: x -transponiert) ist ein K -Vektorraumisomorphismus $K^n \rightarrow (K^n)^*$.

Die duale Paarung $(K^n)^* \times K^n \rightarrow K$ induziert dann das **Standardskalarprodukt**

$$\begin{aligned} \langle , \rangle : K^n \times K^n &\longrightarrow K \\ (x, y) &\longmapsto \langle x, y \rangle := \sum_{i=1}^n x_i y_i = x^t y \end{aligned}$$

auf K^n .

Es gelten die Regeln:

$$\langle x, y \rangle = \langle y, x \rangle \quad (\text{Symmetrie}) \quad (38)$$

$$\left. \begin{aligned} \langle x, b_1 y_1 + b_2 y_2 \rangle &= b_1 \langle x, y_1 \rangle + b_2 \langle x, y_2 \rangle \\ \langle a_1 x_1 + a_2 x_2, y \rangle &= a_1 \langle x_1, y \rangle + a_2 \langle x_2, y \rangle \end{aligned} \right\} \quad (\text{Bilinearitat}) \quad (39)$$

fur alle $x_1, x_2, x, y_1, y_2, y \in K^n, a_1, a_2, b_1, b_2 \in K$. Man sagt: \langle , \rangle ist eine symmetrische Bilinearform auf K^n . \langle , \rangle heit auch die **Standardbilinearform** auf K^n .

Ist

$$A = \begin{pmatrix} a_1 \\ \vdots \\ a_m \end{pmatrix} \in K^{m,n}$$

und $x \in K^n$, so ist nach Definition 11.3

$$Ax = \begin{pmatrix} a_1 x \\ \vdots \\ a_m x \end{pmatrix} = \begin{pmatrix} \langle a_1^t, x \rangle \\ \vdots \\ \langle a_m^t, x \rangle \end{pmatrix},$$

wobei $a_i^t = [a_{i1}, \dots, a_{in}]$ der Spaltenvektor ist, der durch Transponieren der i -ten Zeile von A entsteht.

Definition 11.7 Fur $A = (a_{ij}), B = (b_{ij}) \in K^{m,n}, a \in K$ sei

$$(i) \quad A + B := (a_{ij} + b_{ij}),$$

$$(ii) \quad aA := (aa_{ij}).$$

Dann gilt

Satz 11.8 $K^{m,n}$ ist ein K -Vektorraum, und die kanonische Abbildung

$$\Phi : \text{Hom}(K^n, K^m) \rightarrow K^{m,n}$$

ist ein K -Vektorraumisomorphismus.

Beweis: Es sei $\Phi(\varphi) = A$, $\Phi(\psi) = B$. Dann ist

$$a^i = \varphi(e^i)$$

die i -te Spalte von A und

$$b^i = \psi(e^i)$$

die i -te Spalte von B ; und somit ist

$$(\varphi + \psi)(e^i) = a^i + b^i$$

die i -te Spalte von $A + B$. Also gilt: $\Phi(\varphi + \psi) = A + B$. Genauso folgt $\Phi(a\varphi) = aA$. \square

Definition 11.9 Es sei V ein n -dimensionaler und W ein m -dimensionaler K -Vektorraum.

$\mathcal{A} = (v_1, \dots, v_n)$ sei eine Basis von V , $\mathcal{B} = (w_1, \dots, w_m)$ sei eine Basis von W . $\varphi : V \rightarrow W$ sei linear und $A = (a_{ij}) \in K^{m,n}$ sei die Matrix mit

$$\varphi(v_j) = \sum_{i=1}^m a_{ij} w_i. \quad (40)$$

A heißt **die Matrix von φ bzgl. der Basen \mathcal{A} und \mathcal{B}** . Wir bezeichnen sie auch mit $M_{\mathcal{B}}^{\mathcal{A}}(\varphi)$.

Es gilt nun

Satz 11.10 Es seien $V, W, \mathcal{A}, \mathcal{B}$ wie in Definition 11.9 gegeben. Die Zuordnung $\varphi \mapsto A = (a_{ij})$ mit $\varphi(v_j) = \sum_{i=1}^m a_{ij} w_i$ ist ein K -Vektorraumisomorphismus

$$M_{\mathcal{B}}^{\mathcal{A}} : \text{Hom}(V, W) \rightarrow K^{m,n}.$$

Ist $A = (a_{ij}) \in K^{m,n}$ gegeben, so erhält man φ durch

$$\varphi \left(\sum_{j=1}^n x_j v_j \right) = \sum_{i=1}^m \left(\sum_{j=1}^n a_{ij} x_j \right) w_i.$$

Mit den *Koordinatenabbildungen*

$$\Phi_{\mathcal{A}} : K^n \xrightarrow{\cong} V, \quad \Phi_{\mathcal{B}} : K^m \xrightarrow{\cong} W, \quad \Phi_{\mathcal{A}} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \sum_{j=1}^n x_j v_j, \quad \Phi_{\mathcal{B}} \begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix} = \sum_{i=1}^m y_i w_i$$

erhält man das kommutative Diagramm

$$\begin{array}{ccc} V & \xrightarrow{\varphi} & W \\ \Phi_{\mathcal{A}} \uparrow \cong & & \cong \uparrow \Phi_{\mathcal{B}} \\ K^n & \xrightarrow{(x \mapsto Ax)} & K^m \end{array}$$

mit anderen Worten:

$$M_{\mathcal{B}}^{\mathcal{A}}(\varphi) = (\Phi_{\mathcal{B}})^{-1} \circ \varphi \circ \Phi_{\mathcal{A}}$$

.

Beweis: φ und A bestimmen sich durch (40) gegenseitig, und es gilt:

$$\varphi \circ \Phi_{\mathcal{A}}(x) = \varphi \left(\sum_{j=1}^n x_j v_j \right) = \sum_{i=1}^m \left(\sum_{j=1}^n a_{ij} x_j \right) w_i = \Phi_{\mathcal{B}}(Ax).$$

□

Definition 11.11 Ist $A = (a_{ij}) \in K^{m,n}$, so heißt

$$A^t = (b_{ij}) \in K^{n,m} \quad \text{mit} \quad b_{ij} := a_{ji}$$

die zu A **transponierte Matrix**.

Beispiel 11.12

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

Ist $A = (a^1, \dots, a^n)$, so ist

$$A^t = \begin{pmatrix} a^{1t} \\ \vdots \\ a^{nt} \end{pmatrix},$$

und ebenso ist $A^t = (a_1^t, \dots, a_m^t)$, wenn

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

Satz 11.13 Es sei V ein n -dimensionaler und W ein m -dimensionaler K -Vektorraum.

$\alpha : V \rightarrow W$ sei linear, $\mathcal{A} = (v_1, \dots, v_n)$ sei eine Basis von V , $\mathcal{B} = (w_1, \dots, w_m)$ sei eine Basis von W . $\mathcal{A}^* = (\varphi_1, \dots, \varphi_n)$, $\mathcal{B}^* = (\psi_1, \dots, \psi_m)$ seien die dualen Basen zu \mathcal{A} bzw. \mathcal{B} . Dann gilt:

Ist A die Matrix von α bzgl. \mathcal{A} und \mathcal{B} , so ist A^t die Matrix von α^* bzgl. \mathcal{B}^* und \mathcal{A}^* .

In Formeln:

$$M_{\mathcal{B}}^A(\alpha)^t = M_{\mathcal{A}^*}^{\mathcal{B}^*}(\alpha^*)$$

Beweis: Es sei $A = (a_{ij})$ mit

$$\alpha(v_j) = \sum_{k=1}^m a_{kj} w_k \quad (j = 1, \dots, n),$$

und es sei $B = (b_{ij})$ die Matrix von α^* , also

$$\alpha^*(\psi_i) = \sum_{j=1}^n b_{ji} \varphi_j.$$

Es folgt

$$\begin{aligned}\alpha^*(\psi_i) &= \sum_{j=1}^n \langle \alpha^*(\psi_i) | v_j \rangle \varphi_j = \sum_{j=1}^n \langle \psi_i | \alpha(v_j) \rangle \varphi_j \\ &= \sum_{j=1}^n \sum_{k=1}^m a_{kj} \langle \psi_i | w_k \rangle \varphi_j = \sum_{j=1}^n \sum_{k=1}^m a_{kj} \delta_{ik} \varphi_j \\ &= \sum_{j=1}^m a_{ij} \varphi_j;\end{aligned}$$

und somit $b_{ji} = a_{ij}$, d.h., $B = A^t$. □

Wir wollen jetzt die Matrix $C = (c_{ij}) \in K^{l,n}$ der Komposition $\psi \circ \varphi : K^n \rightarrow K^l$ zweier linearer Abbildungen

$$\varphi : K^n \rightarrow K^m, \quad \psi : K^m \rightarrow K^l$$

bestimmen. Es seien $A \in K^{m,n}$, $B \in K^{l,m}$ mit

$$y = \varphi(x) = Ax, \quad z = \psi(y) = By$$

für

$$x = [x_1, \dots, x_n] \in K^n, \quad y = [y_1, \dots, y_m] \in K^m, \quad z = [z_1, \dots, z_l] \in K^l.$$

Dann gilt also:

$$y_k = \sum_{j=1}^n a_{kj} x_j \quad \text{für } k = 1, \dots, m \quad (41)$$

und

$$z_i = \sum_{k=1}^m b_{ik} y_k \quad \text{für } i = 1, \dots, l. \quad (42)$$

Da auch $z = \psi(y) = \psi \circ \varphi(x) = Cx$ gilt, ist

$$z_i = \sum_{j=1}^n c_{ij} x_j. \quad (43)$$

Setzt man (41) in (42) ein, so erhält man

$$z_i = \sum_{k=1}^m \sum_{j=1}^n b_{ik} a_{kj} x_j = \sum_{j=1}^n \left(\sum_{k=1}^m b_{ik} a_{kj} \right) x_j.$$

Vergleicht man dies mit (43), so ergibt sich damit

$$c_{ij} = \sum_{k=1}^m b_{ik} a_{kj} \quad \text{für } i = 1, \dots, l; \quad j = 1, \dots, n. \quad (44)$$

Wir definieren daher

Definition 11.14 Für $A = (a_{ij}) \in K^{m,n}$, $B \in K^{l,m}$ wird die **Produktmatrix** $BA \in K^{l,n}$ definiert als $BA = (c_{ij})$ mit

$$c_{ij} := \sum_{k=1}^m b_{ik} a_{kj} \quad \text{für } i = 1, \dots, l; j = 1, \dots, n.$$

Wir haben bewiesen:

Satz 11.15 Sind $\varphi : K^n \rightarrow K^m$, $\psi : K^m \rightarrow K^l$ linear mit $\varphi(x) = Ax$, $\psi(y) = By$, so gilt:

$$(\psi \circ \varphi)(x) = (BA)x \quad \text{für alle } x \in K^n.$$

Man beachte, dass nach Definition $(\psi \circ \varphi)(x) = \psi(\varphi(x)) = B(Ax)$ gilt.

Beispiel 11.16 Für

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

ist

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

Offensichtlich gilt auch allgemein

Satz 11.17 Seien $V \xrightarrow{\varphi} W \xrightarrow{\psi} U$ lineare Abbildungen, \mathcal{A} Basis von V , \mathcal{B} Basis von W und \mathcal{C} Basis von U .

A sei die Matrix von φ bzgl. \mathcal{A} und \mathcal{B} ;

B sei die Matrix von ψ bzgl. \mathcal{B} und \mathcal{C} .

Dann ist BA die Matrix von $\psi \circ \varphi$ bzgl. \mathcal{A} und \mathcal{C} .

Beweis: Das Diagramm

$$\begin{array}{ccccc} V & \xrightarrow{\varphi} & W & \xrightarrow{\psi} & U \\ \uparrow \Phi_{\mathcal{A}} & & \uparrow \Phi_{\mathcal{B}} & & \uparrow \Phi_{\mathcal{C}} \\ K^n & \xrightarrow{A} & K^m & \xrightarrow{B} & K^l \end{array}$$

ist kommutativ. Also ist $(\psi \circ \varphi) \Phi_{\mathcal{A}}(x) = \Phi_{\mathcal{C}}((BA)x)$. □

Die Formel (44) kann man auch so lesen:

$$c_{ij} = b_i a^j = \langle b_i^t, a^j \rangle,$$

wobei b_i die i -te Zeile von B und a^j die j -te Spalte von A ist.

$$BA = \begin{pmatrix} \langle b_1^t, a^1 \rangle & \dots & \langle b_1^t, a^n \rangle \\ \vdots & & \vdots \\ \langle b_l^t, a^1 \rangle & \dots & \langle b_l^t, a^n \rangle \end{pmatrix}.$$

$(B, A) \mapsto BA$ ist eine Multiplikation

$$K^{l,m} \times K^{m,n} \longrightarrow K^{l,n} \quad (\text{für alle } l, m, n > 0).$$

Offensichtlich gelten folgende Rechenregeln:

Lemma 11.18 Es sei K ein Körper und $k, l, m, n \in \mathbb{N} \setminus \{0\}$.

- (1) $(A + B)C = AC + BC$ für $A, B \in K^{l,m}$, $C \in K^{m,n}$.
- (2) $A(B + C) = AB + AC$ für $A \in K^{l,m}$, $B, C \in K^{m,n}$.
- (3) $(aA)B = a(AB) = A(aB)$ für $A \in K^{l,m}$, $B \in K^{m,n}$, $a \in K$.
- (4) $(AB)C = A(BC)$ für $A \in K^{k,l}$, $B \in K^{l,m}$, $C \in K^{m,n}$.
- (5) $E_m A = A = A E_n$ für $A \in K^{m,n}$ und
 $E_m = (\delta_{ij})_{i,j=1\dots m} \in K^{m,m}$, $E_n = (\delta_{ij})_{i,j=1\dots n} \in K^{n,n}$.

Beweis: Nach Satz 11.8 und Satz 11.15 kann man $K^{m,n}$ mit $\text{Hom}(K^n, K^m)$ identifizieren.

Der Matrizenmultiplikation entspricht nach Konstruktion die Komposition von linearen Abbildungen.

Der Einheitsmatrix E_m entspricht die Abbildung $id : K^m \rightarrow K^m$. Deshalb gelten die Regeln (1) - (5).

(1) beweist man etwa so: Für $x \in K^n$ ist

$$\begin{aligned} ((A + B)C)x &= (A + B)(Cx) = A(Cx) + B(Cx) \\ &= (AC)x + (BC)x \\ &= (AC + BC)x; \end{aligned}$$

also

$$(A + B)C = AC + BC.$$

□

Lemma 11.19 Für $A \in K^{l,m}$, $B \in K^{m,n}$ gilt:

$$(AB)^t = B^t A^t.$$

Beweis: Sei

$$A = (a_{ij})_{i,j}, \quad B = (b_{jk})_{j,k}.$$

Dann ist

$$AB = \left(\sum_{j=1}^m a_{ij} b_{jk} \right)_{i,k}$$

und

$$(AB)^t = \left(\sum_{j=1}^m b_{jk} a_{ij} \right)_{k,i} = B^t A^t.$$

□

Dieses Lemma folgt natürlich auch aus der Formel $(\psi \circ \varphi)^* = \varphi^* \circ \psi^*$ und den Sätzen 11.13 und 11.17.

Übungen

1. Es sei K ein Körper, $a, b, c \in K$, $m \in \mathbb{N}$. Für $A \in K^{n,n}$ sei A^m die m -te Potenz von A , also $A^0 = E_n$, $A^m = A(A^{m-1})$ für $m > 0$. Beweisen Sie:

$$\begin{pmatrix} 1 & a & c \\ 0 & 1 & b \\ 0 & 0 & 1 \end{pmatrix}^m = \begin{pmatrix} 1 & ma & \binom{m}{2}ab + mc \\ 0 & 1 & mb \\ 0 & 0 & 1 \end{pmatrix}.$$

Welche Bedeutung hat die Matrix auf der rechten Seite für $m = -1$?

2. Es sei

$$\mathbf{L} = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathbb{C}^{2,2} \mid a + d = 0 \right\}.$$

- (a) Zeigen Sie: \mathbf{L} ist ein 3-dimensionaler komplexer Untervektorraum von $\mathbb{C}^{2,2}$. Die Matrizen (**Pauli-Matrizen**)

$$A_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad A_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad A_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

bilden eine \mathbb{C} -Basis von \mathbf{L} . Es gilt: Für $A, B \in \mathbf{L}$ ist $[A, B] := AB - BA \in \mathbf{L}$.

- (b) iA_0, iA_1, iA_2, iA_3 bilden eine \mathbb{C} -Basis von $\mathbb{C}^{2,2}$ und eine \mathbb{R} -Basis des \mathbb{R} -Vektorraumes

$$\mathfrak{su}(2) = \{A \in \mathbb{C}^{2,2} \mid \bar{A} = -A^t\},$$

wobei $A_0 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ und für

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathbb{C}^{2,2}$$

mit \bar{A} die konjugierte Matrix

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

bezeichnet wird.

- (c) Für alle Tripel (a, b, c) mit $\{a, b, c\} = \{1, 2, 3\}$ gibt es ein Vorzeichen $\varepsilon(a, b, c) \in \{\pm 1\}$, so dass

$$[A_a, A_b] = 2i\varepsilon(a, b, c)A_c.$$

Bestimmen Sie $\varepsilon(a, b, c)$ für die sechs möglichen Tripel (a, b, c) .

- (d) Für alle Paare $(a, b) \in \{1, 2, 3\}^2$ gilt:

$$A_a A_b + A_b A_a = 2\delta_{ab} A_0,$$

wobei δ_{ab} das Kroneckersymbol bezeichnet und $A_0 := E_2$ ist.

3. Es seien $M, M' \in K^{4,4}$ als Blockmatrizen

$$M = \begin{pmatrix} A & B \\ C & D \end{pmatrix}, \quad M' = \begin{pmatrix} A' & B' \\ C' & D' \end{pmatrix}$$

mit $A, B, C, D; A', B', C', D' \in K^{2,2}$ gegeben. Wie erhält man MM' ?

4. Mit den Pauli-Matrizen A_0, A_1, A_2, A_3 aus Aufgabe 2 bildet man die Matrizen (**Dirac-Matrizen**)

$$B_0 = \begin{pmatrix} A_0 & 0 \\ 0 & -A_0 \end{pmatrix}, \quad B_1 = \begin{pmatrix} 0 & A_1 \\ -A_1 & 0 \end{pmatrix}, \quad B_2 = \begin{pmatrix} 0 & A_2 \\ -A_2 & 0 \end{pmatrix},$$

$$B_3 = \begin{pmatrix} 0 & A_3 \\ -A_3 & 0 \end{pmatrix}.$$

Beweisen Sie:

- (a) Es sei $\gamma_{ab} = 0$ für $a \neq b$, $\gamma_{00} = 1$, $\gamma_{11} = \gamma_{22} = \gamma_{33} = -1$. Dann gilt für alle $(a, b) \in \{0, 1, 2, 3\}^2$:

$$B_a B_b + B_b B_a = 2\gamma_{ab} E_4.$$

- (b) Berechnen Sie $B := iB_1 B_2 B_3 B_0$ und B^2 .

5. Eine Matrix $P = (p_{ij}) \in \mathbb{R}^{m,n}$ heißt **Markov-Matrix**, wenn $p_{ij} \geq 0$ für alle $i = 1, \dots, m$, $j = 1, \dots, n$ gilt und wenn die Spalten von P **Wahrscheinlichkeitsvektoren** sind, d.h. wenn

$$\sum_{i=1}^m p_{ij} = 1$$

für $j = 1, \dots, n$ gilt.

(a) $P \in \mathbb{R}^{m,n}$ ist genau dann eine Markov-Matrix, wenn gilt:

Für alle Wahrscheinlichkeitsvektoren $x \in \mathbb{R}^n$ ist Px ein Wahrscheinlichkeitsvektor in \mathbb{R}^m .

(b) Sind $P \in \mathbb{R}^{l,m}$, $Q \in \mathbb{R}^{m,n}$ Markov-Matrizen, so ist auch $PQ \in \mathbb{R}^{l,n}$ eine Markov-Matrix. Ist $P = (p_{ij})$ positive Markov-Matrix, d.h. $p_{ij} > 0$ für alle i, j , so ist auch PQ positiv.

(c) Es sei

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

eine positive Markov-Matrix. Zeigen Sie: Es gibt genau eine Markov-Matrix

$$X = \begin{pmatrix} x & y \\ z & w \end{pmatrix} \quad \text{mit } PX = X.$$

Bestimmen Sie X für

$$P = \begin{pmatrix} \frac{1}{2} & \frac{1}{4} \\ \frac{1}{2} & \frac{3}{4} \end{pmatrix},$$

und vergleichen Sie X mit P^4 .

12 Der Rang einer Matrix: Äquivalenz von Matrizen

Es sei K ein Körper.

Definition 12.1 Es sei $A \in K^{m,n}$ eine $m \times n$ -Matrix mit den Zeilen a_1, \dots, a_m und den Spalten a^1, \dots, a^n .

$$\text{Z-rg}(A) = \dim\langle a_1, \dots, a_m \rangle$$

heißt der **Zeilenrang** von A .

$$\text{S-rg}(A) = \dim\langle a^1, \dots, a^n \rangle$$

heißt der **Spaltenrang** von A .

Satz 12.2 Ist $\varphi : V \rightarrow W$ linear, $\mathcal{A} = (v_1, \dots, v_n)$ Basis von V , $\mathcal{B} = (w_1, \dots, w_m)$ Basis von W , $A \in K^{m,n}$ die Matrix von φ bzgl. \mathcal{A} und \mathcal{B} , so gilt:

$$\text{rg}(\varphi) = \text{S-rg}(A)$$

$$\text{rg}(\varphi^*) = \text{S-rg}(A^t) = \text{Z-rg}(A).$$

Beweis: Es gilt:

$$\varphi(\Phi_{\mathcal{A}}(x)) = \Phi_{\mathcal{B}}(Ax) \quad \text{für alle } x \in K^n.$$

Für die Basisvektoren e^j von K^n gilt:

$$\Phi_{\mathcal{A}}(e^j) = v_j \quad \text{und} \quad Ae^j = a^j,$$

also

$$\varphi(v_j) = \Phi_{\mathcal{B}}(a^j);$$

und somit ist

$$\text{im } \varphi = \langle \varphi(v_1), \dots, \varphi(v_n) \rangle = \Phi_{\mathcal{B}}(\langle a^1, \dots, a^n \rangle).$$

Da $\Phi_{\mathcal{B}} : K^m \rightarrow W$ ein Isomorphismus ist, gilt:

$$\text{rg}(\varphi) = \dim(\text{im } \varphi) = \dim\langle a^1, \dots, a^n \rangle = \text{S-rg}(A).$$

Da φ^* die Matrix A^t besitzt, gilt auch:

$$\text{rg}(\varphi^*) = \text{S-rg}(A^t) = \text{Z-rg}(A).$$

□

Hieraus und aus dem Satz $\text{rg}(\varphi) = \text{rg}(\varphi^*)$ folgt nun

Satz 12.3 Für jede Matrix $A \in K^{m,n}$ gilt:

$$\text{Z-rg}(A) = \text{S-rg}(A).$$

Diese Zahl nennen wir jetzt kurz den **Rang** von A und bezeichnen sie mit $\text{rg}(A)$.
Für $A \in K^{m,n}$ ist $\text{rg}(A) \leq \min(m, n)$.

Definition 12.4 $A \in K^{m,n}$ hat **maximalen Rang**, falls $\text{rg}(A) = \min(m, n)$.

Ist $m \leq n$, so heißt das: Die Zeilen von A sind linear unabhängig.

Ist $m \geq n$, so bedeutet das: Die Spalten von A sind linear unabhängig.

Satz 12.5 Ist $\varphi : V \rightarrow W$ linear mit Matrix $A \in K^{m,n}$, so gilt: A hat maximalen Rang genau dann, wenn φ injektiv oder surjektiv ist.

Beweis: 1. Fall:

$$\dim V = n \leq m = \dim W.$$

Dann gilt: φ injektiv $\Leftrightarrow \ker \varphi = 0 \Leftrightarrow \text{rg}(\varphi) = n \Leftrightarrow \text{rg}A = n \Leftrightarrow$ Die Spalten a^1, \dots, a^n sind linear unabhängig.

2. Fall:

$$n \geq m.$$

Dann gilt: φ ist surjektiv $\Leftrightarrow \text{im } \varphi = W \Leftrightarrow \text{rg}(\varphi) = m \Leftrightarrow \text{rg}A = m \Leftrightarrow$ Die Zeilen a_1, \dots, a_m sind linear unabhängig. \square

Definition 12.6 Eine Matrix $A \in K^{n,n}$ heißt **invertierbar**, wenn es eine Matrix $B \in K^{n,n}$ mit

$$AB = BA = E_n$$

gibt. Ist A invertierbar, so ist B eindeutig durch A bestimmt und wird mit A^{-1} bezeichnet. A^{-1} heißt die **inverse Matrix** von A .

Satz 12.7 Sind V, W n -dimensionale K -Vektorräume und ist $\mathcal{A} = (v_1, \dots, v_n)$ Basis von V , $\mathcal{B} = (w_1, \dots, w_n)$ Basis von W , $\varphi : V \rightarrow W$ linear, $A \in K^{n,n}$ die Matrix von φ bzgl. \mathcal{A} und \mathcal{B} , so gilt:

A ist invertierbar $\Leftrightarrow \varphi$ ist Isomorphismus.

Beweis: " \Rightarrow ": Ist $AB = BA = E_n$, so definiert B eine lineare Abbildung $\psi : W \rightarrow V$ mit

$$\psi(\Phi_{\mathcal{B}}(y)) = \Phi_{\mathcal{A}}(By) \quad \forall y \in K^n.$$

Es gilt außerdem:

$$\varphi(\Phi_{\mathcal{A}}(x)) = \Phi_{\mathcal{B}}(Ax) \quad \forall x \in K^n.$$

Es folgt

$$\psi(\varphi(\Phi_{\mathcal{A}}(x))) = \psi(\Phi_{\mathcal{B}}(Ax)) = \Phi_{\mathcal{A}}(BAx) = \Phi_{\mathcal{A}}(x),$$

also $\psi \circ \varphi = id_V$; und analog

$$\varphi(\psi(\Phi_{\mathcal{B}}(y))) = \varphi(\Phi_{\mathcal{A}}(By)) = \Phi_{\mathcal{B}}(ABY) = \Phi_{\mathcal{B}}(y),$$

also $\varphi \circ \psi = id_W$.

“ \Leftarrow ” Ist $\psi \circ \varphi = id_V$, $\varphi \circ \psi = id_W$ und B die Matrix von ψ bzgl. \mathcal{B} und \mathcal{A} , so ist AB die Matrix von $\varphi \circ \psi$ bzgl. \mathcal{B} und \mathcal{B} , und da E_n die Matrix von id_W bzgl. \mathcal{B} und \mathcal{B} ist, folgt $AB = E_n$. Genauso folgt $BA = E_n$. \square

Da für eine lineare Abbildung $\varphi : V \rightarrow W$ zwischen zwei n -dimensionalen K -Vektorräumen die Aussagen

- φ ist Isomorphismus,
- φ ist injektiv,
- φ ist surjektiv

äquivalent sind, ergibt sich aus 12.5 und 12.7 die Folgerung

Satz 12.8 Für eine Matrix $A \in K^{n,n}$ sind äquivalent:

- a) A ist invertierbar.
- b) $\text{rg}(A) = n$.
- c) Die Zeilen von A sind linear unabhängig.
- d) Die Spalten von A sind linear unabhängig.

Definition 12.9 Es sei X irgendeine nicht leere Menge, und es sei $A \in X^{m,n}$ eine $m \times n$ -Matrix. Es seien

$$\mathbf{i} = (i_1, \dots, i_p) \in \{1, \dots, m\}^p$$

und

$$\mathbf{j} = (j_1, \dots, j_q) \in \{1, \dots, n\}^q$$

beliebig gewählte Multiindizes. Dann wird

$$A_{\mathbf{ij}}$$

definiert als die $p \times q$ -Matrix

$$B = (b_{\nu\mu}) \in X^{p,q}$$

mit

$$b_{\nu\mu} := a_{i_\nu j_\mu}.$$

Ist $i_1 < \dots < i_p$ und $j_1 < \dots < j_q$, so ist $A_{\mathbf{ij}}$ die $p \times q$ -**Untermatrix** von A aus den Elementen in den Zeilen i_1, \dots, i_p und den Spalten j_1, \dots, j_q . Man schreibt ausführlich

$$A_{\mathbf{ij}} = A_{(i_1, \dots, i_p)(j_1, \dots, j_q)}.$$

Beispiel 12.10 Sei

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 0 & 7 & 0 \end{pmatrix}.$$

Dann ist zum Beispiel

$$A_{11} = 1, \quad A_{(1,2)(1,2)} = \begin{pmatrix} 1 & 2 \\ 4 & 5 \end{pmatrix}, \quad A_{(1,1)(1,1)} = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}, \quad A_{(2,1)(1,1)} = \begin{pmatrix} 4 & 4 \\ 1 & 1 \end{pmatrix},$$

$$A_{(1,2)(2,3)} = \begin{pmatrix} 2 & 3 \\ 5 & 6 \end{pmatrix}, \quad A_{(1,2,3)(1,2,3)} = A, \quad A_{2(1,2,3)} = (4 \ 5 \ 6).$$

Mann kann A auch als Blockmatrix schreiben, z.B.

$$A = \left(\begin{array}{c|c} A_{(1,2)(1,2)} & A_{(1,2)3} \\ \hline A_{3(1,2)} & A_{33} \end{array} \right).$$

Satz 12.11 Es sei $A \in K^{m,n}$ und $r \in \mathbb{N}$. Dann sind äquivalent:

- (i) $\text{rg}(A) \geq r$.
- (ii) Es gibt eine invertierbare $r \times r$ -Untermatrix von A , d.h. $\exists \mathbf{i} = (i_1, \dots, i_r)$, $\mathbf{j} = (j_1, \dots, j_r)$ mit $1 \leq i_1 < \dots < i_r \leq m$, $1 \leq j_1 < \dots < j_r \leq n$, so dass

$$A_{\mathbf{ij}} \in K^{r,r}$$

invertierbar ist.

Beweis: (i) \Rightarrow (ii):

Ist $\text{rg}(A) \geq r$, so sind r Spalten a^{j_1}, \dots, a^{j_r} von A linear unabhängig. Die Matrix $A_{(1,\dots,m)(j_1,\dots,j_r)}$ hat den Rang r . Also sind r Zeilen dieser Matrix linear unabhängig, etwa die Zeilen i_1, \dots, i_r . Dann ist die Untermatrix

$$A_{(i_1,\dots,i_r)(j_1,\dots,j_r)}$$

invertierbar.

(ii) \Rightarrow (i):

Ist $A_{\mathbf{ij}}$ invertierbar, so sind die Zeilen i_1, \dots, i_r von $A_{(1,\dots,m)\mathbf{j}}$ linear unabhängig, also sind auch die Spalten von $A_{(1,\dots,m)\mathbf{j}}$ linear unabhängig, also ist $\text{rg}A \geq r$. \square

Beispiel 12.12 Es sei

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

Dann ist $\text{rg}A \geq 2$, weil $\begin{pmatrix} 1 & 3 \\ 0 & 1 \end{pmatrix}$ invertierbar ist:

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

Es gilt natürlich $\text{rg}A = 3$, wie man durch elementare Zeilenumformungen sieht, die den Zeilenrang der Matrix natürlich nicht ändern, weil sich der von den Zeilen aufgespannte Untervektorraum nicht ändert:

$$A = \begin{pmatrix} 1 & 3 & 1 \\ 0 & 1 & 0 \\ 4 & 1 & 2 \end{pmatrix} \xrightarrow{T_{31}(-4)} \begin{pmatrix} 1 & 3 & 1 \\ 0 & 1 & 0 \\ 0 & -11 & -2 \end{pmatrix} \xrightarrow{T_{32}(11)} \begin{pmatrix} 1 & 3 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{pmatrix}.$$

Die Zeilen der letzten Matrix sind linear unabhängig, also sind es auch die Zeilen von A , d.h. $\text{rg}A = 3$. A ist also auch invertierbar.

Wir wollen demonstrieren, wie man die inverse Matrix A^{-1} berechnen kann.

Man bringt A durch Zeilenumformungen auf die reduzierte Zeilenstufenform, welche hier automatisch die Einheitsmatrix ist. Führt man die Zeilenumformungen simultan auch an der Einheitsmatrix durch, so wird diese gerade in A^{-1} transformiert! In unserem Beispiel ergibt sich:

$$\begin{aligned}
 A &= \begin{pmatrix} 1 & 3 & 1 \\ 0 & 1 & 0 \\ 4 & 1 & 2 \end{pmatrix} & E_3 &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \\
 T_{31}(-4) : & \begin{pmatrix} 1 & 3 & 1 \\ 0 & 1 & 0 \\ 0 & -11 & -2 \end{pmatrix} & & \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -4 & 0 & 1 \end{pmatrix} \\
 T_{32}(11) : & \begin{pmatrix} 1 & 3 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{pmatrix} & & \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -4 & 11 & 1 \end{pmatrix} \\
 T_{12}(-3) : & \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{pmatrix} & & \begin{pmatrix} 1 & -3 & 0 \\ 0 & 1 & 0 \\ -4 & 11 & 1 \end{pmatrix} \\
 T_3\left(-\frac{1}{2}\right) : & \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} & & \begin{pmatrix} 1 & -3 & 0 \\ 0 & 1 & 0 \\ +2 & -\frac{11}{2} & -\frac{1}{2} \end{pmatrix} \\
 T_{13}(-1) : & \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} & & \begin{pmatrix} -1 & +\frac{5}{2} & \frac{1}{2} \\ 0 & 1 & 0 \\ 2 & -\frac{11}{2} & -\frac{1}{2} \end{pmatrix} =: B.
 \end{aligned}$$

Man errechnet:

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

Also ist $A^{-1} = B$. Wieso geht das?

Der Grund ist folgender:

Die Zeilenumformung $T_{31}(-4)$ auf A anwenden, bedeutet: A von links mit der Matrix

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -4 & 0 & 1 \end{pmatrix}$$

zu multiplizieren:

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -4 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 3 & 1 \\ 0 & 1 & 0 \\ 4 & 1 & 2 \end{pmatrix} = \begin{pmatrix} 1 & 3 & 1 \\ 0 & 1 & 0 \\ 0 & -11 & -2 \end{pmatrix} \quad \text{usw.}$$

Allgemein gilt

Lemma 12.13 Es sei $A \in K^{m,n}$, und es seien $a \in K$, $i, j \in \{1, \dots, m\}$, $i \neq j$. Dann gilt:

- a) Wendet man die elementare Zeilenoperation $T_{ij}(a)$ auf die Zeilen der Matrix A an, so ergibt sich die Matrix

$$A_1 = Q_{ij}(a)A,$$

wobei

$$Q_{ij}(a) = Q_{ij}^{(m)}(a) \in K^{m,m}$$

die **Elementarmatrix**

$$Q_{ij}(a) = E_m + aE_{ij}^{(m)} = \begin{pmatrix} 1 & \dots & 0 & \dots & 0 & \dots & 0 \\ \vdots & \ddots & \vdots & & \vdots & & \vdots \\ 0 & \dots & 1 & \dots & 0 & \dots & 0 \\ \vdots & & \vdots & \ddots & \vdots & & \vdots \\ 0 & \dots & a & \dots & 1 & \dots & 0 \\ \vdots & & \vdots & & \vdots & \ddots & \vdots \\ 0 & \dots & 0 & \dots & 0 & \dots & 1 \end{pmatrix}$$

ist und $E_{ij}^{(m)} := (e_{\nu\mu})$ mit

$$e_{\nu\mu} = \begin{cases} 1 & \text{falls } (\nu, \mu) = (i, j) \\ 0 & \text{sonst.} \end{cases}$$

- b) Die elementare Zeilenoperation T_{ij} auf A anzuwenden, bedeutet, A von links mit der **Elementarmatrix**

$$P_{ij} = E_m - E_{ii} - E_{jj} + E_{ij} + E_{ji} = \begin{pmatrix} 1 & & & & & & & 0 \\ & \ddots & & & & & & \\ & & 0 & \dots & 1 & & & \\ & & \vdots & \ddots & \vdots & & & \\ & & 1 & \dots & 0 & & & \\ & & & & & \ddots & & \\ 0 & & & & & & & 1 \end{pmatrix}$$

zu multiplizieren.

- c) Die elementare Zeilenoperation $T_i(a)$ ($a \neq 0$) entspricht einer Multiplikation von links mit der **Elementarmatrix**

$$S_i(a) = E_m + (a - 1)E_{ii} = \begin{pmatrix} 1 & & & & 0 \\ & \ddots & & & \\ & & a & & \\ & & & \ddots & \\ 0 & & & & 1 \end{pmatrix}.$$

(a steht in der i -ten Zeile.)

Im obigen Beispiel ist

$$Q_{13}(-1) S_3\left(-\frac{1}{2}\right) Q_{12}(-3) Q_{32}(11) Q_{31}(-4) A = E_3;$$

also

$$A^{-1} = Q_{13}(-1) S_3\left(-\frac{1}{2}\right) Q_{12}(-3) Q_{32}(11) Q_{31}(-4).$$

Damit liefert das Gaußsche Eliminationsverfahren auch einen Algorithmus zur Bestimmung der inversen Matrix einer beliebigen quadratischen Matrix $A \in K^{n,n}$: Durch sukzessive Linksmultiplikation mit Elementarmatrizen B_1, \dots, B_N kann man A auf die reduzierte Zeilenstufenform

$$B_N B_{N-1} \dots B_2 B_1 A = B = \left(\begin{array}{cccccccc} 0 & \dots & 1 & * \dots * & 0 & * \dots * & 0 & \dots & 0 \\ & & & & 1 & * \dots * & 0 & \dots & 0 \\ & & & & & & 1 & \dots & 0 \\ & & & & & & & \ddots & \vdots \\ & & & & & & & & 1 & * \dots \\ & & & & & & & & & 0 \end{array} \right) \left. \vphantom{\begin{pmatrix} 0 \\ \dots \\ 1 \\ \dots \\ 1 \\ \dots \\ 0 \end{pmatrix}} \right\} r \text{ Zeilen}$$

bringen. Es gilt dann: $r = \text{rg}(A)$, und A ist genau dann invertierbar, wenn $r = n$ gilt, und das ist genau dann der Fall, wenn $B = E_n$ ist.

Es gilt dann:

$$A^{-1} = B_N B_{N-1} \dots B_2 B_1.$$

A^{-1} ist also ein Produkt von Elementarmatrizen der Form $Q_{ij}(a)$, $i \neq j$, $a \in K$, P_{ij} , $S_i(a)$, $a \in K \setminus \{0\}$. Es folgt

$$\begin{aligned} A &= (B_N B_{N-1} \dots B_2 B_1)^{-1} \\ &= B_1^{-1} B_2^{-1} \dots B_{N-1}^{-1} B_N^{-1}. \end{aligned}$$

Nun gilt aber:

$$\begin{aligned} Q_{ij}(a)^{-1} &= Q_{ij}(-a), \\ S_i(a)^{-1} &= S_i(a^{-1}), \\ P_{ij}^{-1} &= P_{ij}; \end{aligned}$$

d.h. das Inverse einer Elementarmatrix ist wieder eine Elementarmatrix. Also ist auch A Produkt von Elementarmatrizen. Wir haben damit

Satz 12.14 Jede invertierbare Matrix $A \in K^{n,n}$ ist ein Produkt von Elementarmatrizen.

Definition 12.15 Zwei Matrizen $A, B \in K^{m,n}$ heißen **äquivalent** ($A \sim B$), wenn es invertierbare Matrizen $S \in K^{m,m}$, $T \in K^{n,n}$ gibt, so dass

$$B = SAT$$

gilt.

Dadurch ist eine Äquivalenzrelation auf $K^{m,n}$ definiert, denn:

- a) $A \sim A$, weil $A = E_m A E_n$.
- b) $A \sim B \Rightarrow B = SAT \Rightarrow A = S^{-1} B T^{-1} \Rightarrow B \sim A$.
- c) $A \sim B$ und $B \sim C \Rightarrow B = SAT, C = S' B T' \Rightarrow C = (S'S) A (T T') \Rightarrow A \sim C$.

Welches sind die Äquivalenzklassen?

Satz 12.16 a) Für jede Matrix $A \in K^{m,n}$ mit $r = \text{rg}(A)$ gilt: Es gibt invertierbare Matrizen $S \in K^{m,m}$, $T \in K^{n,n}$, so dass

$$SAT = \left(\begin{array}{c|c} E_r & 0 \\ \hline 0 & \end{array} \right).$$

- b) Zwei Matrizen $A, B \in K^{m,n}$ sind genau dann äquivalent, wenn $\text{rg}(A) = \text{rg}(B)$ gilt.

Beweis: zu a): Nach dem Gaußschen Eliminationsverfahren gibt es eine invertierbare Matrix $S \in K^{m,m}$, so dass

$$SA = \left(\begin{array}{cccc} 0 & \dots & b_{1j_1} & * \dots \\ & & 0 & b_{2j_2} & * \dots \\ & & & & \ddots \\ & & & & 0 & b_{rj_r} & * \dots \\ & & & & & & 0 \end{array} \right)$$

mit $1 \leq j_1 < \dots < j_r \leq n$ und $b_{1j_1} = \dots = b_{rj_r} = 1$.

Jetzt betrachte man die transponierte Matrix

$$(SA)^t = \left(\begin{array}{cccc} 0 & 0 & \dots & 0 \\ \vdots & & & \\ b_{1j_1} & & & \\ * & 0 & & 0 \\ & b_{2j_2} & & 0 \\ \vdots & & \ddots & \\ & * & & b_{rj_r} \\ & \vdots & & \\ & & & * \\ & & & \vdots \end{array} \right).$$

Nach dem Gaußschen Eliminationsverfahren gibt es jetzt eine invertierbare Matrix $T' \in K^{n,n}$, so dass

$$T'(SA)^t = \left(\begin{array}{c|c} E_r & 0 \\ \hline 0 & 0 \end{array} \right).$$

Setzt man $T := T'^t$, so erhält man nach Transponieren von $T'(SA)^t$ die gewünschte Form

$$SAT = \left(\begin{array}{c|c} E_r & 0 \\ \hline 0 & 0 \end{array} \right).$$

Zu b): Sind $A, B \in K^{m,n}$ äquivalent, so gilt offensichtlich: $\text{rg}A = \text{rg}B$. Ist umgekehrt $r = \text{rg}A = \text{rg}B$, so folgt aus a), dass

$$A \sim \left(\begin{array}{c|c} E_r & 0 \\ \hline 0 & 0 \end{array} \right) \sim B.$$

□

Der Beweis liefert ein Rechenverfahren zur Bestimmung der Matrizen S und T . Zunächst bemerken wir:

a)

$$\begin{aligned} AQ_{ij}(a) &= \begin{pmatrix} a_1 \\ \vdots \\ a_m \end{pmatrix} (e^1, \dots, e^i, \dots, e^j + ae^i, \dots, e^n) \\ & \qquad \qquad \qquad \uparrow \\ & \qquad \qquad \qquad j\text{-te Stelle} \\ &= (a^1, \dots, a^i, \dots, a^j + aa^i, \dots, a^n), \\ & \qquad \qquad \qquad \uparrow \\ & \qquad \qquad \qquad j\text{-te Stelle} \end{aligned}$$

d.h. $AQ_{ij}(a)$ entsteht aus A durch Addition des a -fachen der i -ten Spalte zur j -ten Spalte.

b)

$$\begin{aligned} AP_{ij} &= \begin{pmatrix} a_1 \\ \vdots \\ a_m \end{pmatrix} (e^1, \dots, e^j, \dots, e^i, \dots, e^n) \\ & \qquad \qquad \qquad \uparrow \quad \quad \uparrow \\ & \qquad \qquad \qquad i\text{-te Stelle} \quad j\text{-te Stelle} \\ &= (a^1, \dots, a^j, \dots, a^i, \dots, a^n). \\ & \qquad \qquad \qquad \uparrow \quad \quad \uparrow \\ & \qquad \qquad \qquad i\text{-te Stelle} \quad j\text{-te Stelle} \end{aligned}$$

(Vertauschen von Spalten)

c)

$$AS_i(a) = (a^1, \dots, aa^i, \dots, a^n).$$

Wir erhalten nun die Normalform einschließlich der Transformationsmatrizen S, T durch folgendes Verfahren:

Es sei $A \in K^{m,n}$. Die Zeilenoperationen werden simultan an A und an der Einheitsmatrix E_m ausgeführt und die Spaltenoperationen simultan an A und der Einheitsmatrix E_n .

$$\begin{array}{c|c|c}
 \text{Zeilenoperationen} & & \text{Spaltenoperationen} \\
 \hline
 E_m & A & E_n \\
 S_1 E_m & S_1 A T_1 & E_n T_1 \\
 S_2 S_1 E_m & S_2 S_1 A T_1 & \\
 \vdots & \vdots & \vdots \\
 S = S_N \dots S_1 E_m & SAT = \begin{pmatrix} E_r & 0 \\ 0 & 0 \end{pmatrix} & E_n T_1 T_2 \dots T_M = T
 \end{array}$$

Beispiel 12.17

$$\begin{array}{ccc}
 \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} & A = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 4 \end{pmatrix} & \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \\
 \\
 S = \begin{pmatrix} 1 & 0 \\ -2 & 1 \end{pmatrix} & \begin{pmatrix} 1 & 2 & 3 \\ 0 & -3 & -2 \end{pmatrix} & \\
 \\
 & \begin{pmatrix} 1 & 0 & 0 \\ 0 & -3 & -2 \end{pmatrix} & \begin{pmatrix} 1 & -2 & -3 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \\
 \\
 & \begin{pmatrix} 1 & 0 & 0 \\ 0 & -3 & 1 \end{pmatrix} & \begin{pmatrix} 1 & -2 & \frac{3}{2} \\ 0 & 1 & 0 \\ 0 & 0 & -\frac{1}{2} \end{pmatrix} \\
 \\
 SAT = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} & \begin{pmatrix} 1 & \frac{3}{2} & \frac{5}{2} \\ 0 & 0 & 1 \\ 0 & -\frac{1}{2} & -\frac{3}{2} \end{pmatrix} = T
 \end{array}$$

S und T sind nicht eindeutig bestimmt. Man hätte zum Beispiel nach dem ersten Schritt folgendermaßen fortfahren können:

$$\begin{array}{ccc}
 \begin{pmatrix} 1 & 0 \\ \frac{2}{3} & -\frac{1}{3} \end{pmatrix} & \begin{pmatrix} 1 & 2 & 3 \\ 0 & 1 & \frac{2}{3} \end{pmatrix} & \\
 \\
 S = \begin{pmatrix} -\frac{1}{3} & +\frac{2}{3} \\ \frac{2}{3} & -\frac{1}{3} \end{pmatrix} & \begin{pmatrix} 1 & 0 & \frac{5}{3} \\ 0 & 1 & \frac{2}{3} \end{pmatrix}; & T = \begin{pmatrix} 1 & 0 & -\frac{5}{3} \\ 0 & 1 & -\frac{2}{3} \\ 0 & 0 & 1 \end{pmatrix}.
 \end{array}$$

Wir wollen jetzt zeigen, dass zwei Matrizen A und B aus $K^{m,n}$ genau dann äquivalent sind, wenn sie Matrizen ein und derselben linearen Abbildung $\varphi : V \rightarrow W$ bzgl. verschiedener Basen sind.

Sei also $\varphi : V \rightarrow W$ linear:

$$\begin{aligned} \mathcal{A} &= (v_1, \dots, v_n), & \mathcal{A}' &= (v'_1, \dots, v'_n) & \text{seien Basen von } V; \\ \mathcal{B} &= (w_1, \dots, w_m), & \mathcal{B}' &= (w'_1, \dots, w'_m) & \text{seien Basen von } W. \end{aligned}$$

Es sei A die Matrix von φ bzgl. \mathcal{A} und \mathcal{B} und B die Matrix von φ bzgl. \mathcal{A}' und \mathcal{B}' ; also

$$\Phi_{\mathcal{B}}(Ax) = \varphi(\Phi_{\mathcal{A}}(x))$$

und

$$\Phi_{\mathcal{B}'}(Bx) = \varphi(\Phi_{\mathcal{A}'}(x))$$

für alle $x \in K^n$. Dann ist also

$$\begin{aligned} Bx &= (\Phi_{\mathcal{B}'}^{-1} \circ \varphi \circ \Phi_{\mathcal{A}'})(x) & \text{und} \\ \Phi_{\mathcal{B}'}^{-1} \Phi_{\mathcal{B}}(Ax) &= \Phi_{\mathcal{B}'}^{-1} \circ \varphi \circ \Phi_{\mathcal{A}}(x) \\ &= \Phi_{\mathcal{B}'}^{-1} \circ \varphi \circ \Phi_{\mathcal{A}'}(\Phi_{\mathcal{A}'}^{-1} \Phi_{\mathcal{A}})(x) \\ &= B(\Phi_{\mathcal{A}'}^{-1} \Phi_{\mathcal{A}}(x)). \end{aligned}$$

Setzt man

$$S = \Phi_{\mathcal{B}'}^{-1} \Phi_{\mathcal{B}} : K^m \rightarrow K^m$$

und

$$R = \Phi_{\mathcal{A}'}^{-1} \Phi_{\mathcal{A}} : K^n \rightarrow K^n,$$

so erhält man also

$$SA = BR$$

und somit

$$B = SAR^{-1},$$

d.h. $B \sim A$.

Weiter sehen wir: Ist $S = (s_{ij})$, so gilt für $y \in K^m$

$$\Phi_{\mathcal{B}'}(Sy) = \Phi_{\mathcal{B}}(y),$$

also gilt für $y' = Sy$:

$$\sum_{i=1}^m y'_i w'_i = \sum_{i=1}^m \left(\sum_{j=1}^m s_{ij} y_j \right) w'_i = \sum_{i=1}^m y_i w_i,$$

d.h. die Matrix S ist die Transformationsmatrix, welche die Koordinaten y von $w \in W$ bzgl. \mathcal{B} in die Koordinaten y' von $w \in W$ bzgl. \mathcal{B}' transformiert:

$$y'_i = \sum_{j=1}^m s_{ij} y_j.$$

Ebenso ist $R = (r_{ij}) \in K^{n,n}$ die Transformationsmatrix, die die Koordinaten x von $v \in V$ bzgl. \mathcal{A} in die Koordinaten x' von $v \in V$ bzgl. \mathcal{A}' transformiert:

$$x'_i = \sum_{j=1}^n r_{ij} x_j.$$

Kurz zusammengefasst: Ist $y = Ax$, $y' = Bx'$, $y' = Sy$, $x' = Rx$ so folgt $BRx = Bx' = y' = Sy = SAx$ und somit $BR = SA$ also $B = SAR^{-1}$.

Der folgende Satz sagt, dass man zu einer linearen Abbildung zwischen endlichdimensionalen Vektorräumen immer Basen dieser Vektorräume finden kann, bezüglich derer die lineare Abbildung eine besonders einfache Matrix hat, nämlich

$$\begin{pmatrix} E_r & 0 \\ 0 & 0 \end{pmatrix}.$$

Satz 12.18 Ist $\varphi : V \rightarrow W$ linear und $\text{rg}(\varphi) = r$, so gibt es eine Basis $\mathcal{A}' = (v'_1, \dots, v'_n)$ von V und eine Basis $\mathcal{B}' = (w'_1, \dots, w'_m)$ von W , so dass gilt:

$$\varphi(v'_i) = \begin{cases} w'_i & \text{für } i = 1, \dots, r \\ 0 & \text{für } i = r + 1, \dots, n \end{cases}.$$

Beweis: Es seien $\mathcal{A} = (v_1, \dots, v_n)$, $\mathcal{B} = (w_1, \dots, w_m)$ irgendwelche Basen von V bzw. W . Sei $A \in K^{m,n}$ die Matrix von φ bezüglich \mathcal{A} und \mathcal{B} . Dann gibt es invertierbare Matrizen S und T , so dass

$$SAT = B = \begin{pmatrix} E_r & 0 \\ 0 & 0 \end{pmatrix}.$$

Es sei $\mathcal{A}' = (v'_1, \dots, v'_n)$ die Basis von V mit $v'_i = \Phi_{\mathcal{A}}(Te^i)$, also $\Phi_{\mathcal{A}'} = \Phi_{\mathcal{A}} \circ T$. Weiter sei $\mathcal{B}' = (w'_1, \dots, w'_m)$ die Basis von W mit $w'_j = \Phi_{\mathcal{B}}(S^{-1}e^j)$, also $\Phi_{\mathcal{B}'} = \Phi_{\mathcal{B}} \circ S$.

Es folgt nun leicht, dass B die Matrix von φ bezüglich der Basen \mathcal{A}' und \mathcal{B}' ist, denn:

$$\varphi(\Phi_{\mathcal{A}'}(x')) = \varphi(\Phi_{\mathcal{A}}(Tx')) = \Phi_{\mathcal{B}}(ATx') = \Phi_{\mathcal{B}'}(SATx').$$

□

Wir wollen zum Schluss ein Rechenverfahren zur Bestimmung der Matrix B einer linearen Abbildung $A : K^n \rightarrow K^m$ bezüglich zweier Basen $\mathcal{A} = (s^1, \dots, s^n)$ und $\mathcal{B} = (t^1, \dots, t^m)$ von K^n bzw. K^m erläutern. Dabei wird A wegen der kanonischen Identifizierung $\text{Hom}(K^n, K^m) = K^{m,n}$ wie immer als Matrix aufgefasst (die Matrix der linearen Abbildung bezüglich der Standardbasen). Die Basisvektoren in \mathcal{A} bilden die Spalten einer $n \times n$ -Matrix $S = (s^1, \dots, s^n)$ und die Basisvektoren in \mathcal{B} bilden die Spalten einer $m \times m$ -Matrix $R = (r^1, \dots, r^m)$. Für die Koeffizienten b_{ij} von B gilt dann $As^i = \sum_{j=1}^m b_{ji} r^j = Rb^i$, wobei b^i die i -te Spalte von B bezeichnet. Fasst man dies zusammen, so ergibt sich $AS = RB$, also $B = R^{-1}AS$. Um B zu bestimmen

kann man nun etwa direkt die nach Voraussetzung eindeutig lösbaren inhomogenen linearen Gleichungssysteme

$$Rb^i = As^i$$

für die unbekanntenen Spaltenvektoren b^i lösen. Man kann aber auch zunächst AS berechnen und dann durch simultane Zeilenoperationen an AS und R die Matrix R in die Einheitsmatrix transformieren und demzufolge AS in die gesuchte Matrix $B = R^{-1}AS$.

Übungen

1. Beweisen Sie:

$$\operatorname{rg}(AB) \leq \min(\operatorname{rg}(A), \operatorname{rg}(B))$$

für $A \in K^{l,m}$, $B \in K^{m,n}$.

2. Bestimmen Sie den Rang von

$$A = \begin{pmatrix} 1 & 4 & 1 \\ 4 & 1 & 4 \end{pmatrix} \in \mathbb{Z}^{2,3}$$

über \mathbb{Q} , und finden Sie invertierbare Matrizen $S \in \mathbb{Q}^{2,2}$, $R \in \mathbb{Q}^{3,3}$, so dass

$$SAR^{-1} = \begin{pmatrix} E_r & 0 \\ 0 & 0 \end{pmatrix}$$

gilt. Untersuchen Sie dieselbe Frage über \mathbb{F}_p für die Matrix $A \bmod p$, p Primzahl.

3. Ist $A \in K^{m,n}$ und $r = \operatorname{rg}A$, so gibt es Matrizen $B \in K^{m,r}$, $C \in K^{r,n}$, so dass

$$A = BC.$$

Behandeln Sie die Beispiele

$$A = \begin{pmatrix} 3 & 6 & 9 \\ 2 & 4 & 6 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 \\ 2 & 4 & 7 \end{pmatrix}, \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 1 & 2 & 2 \end{pmatrix}.$$

4. Für

$$A = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix} \in \mathbb{F}_2^{4,3}$$

bestimme man invertierbare Matrizen S und T , so dass

$$SAT = \begin{pmatrix} E_r & 0 \\ 0 & 0 \end{pmatrix}$$

gilt.

5. Bestimmen Sie die Matrix von $A : K^2 \rightarrow K^3$, $A \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x + 3y \\ 3x - y \\ 2x + y \end{pmatrix}$ bezüglich der Basen $\mathcal{A} = (v^1, v^2)$, $v^1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$, $v^2 = \begin{pmatrix} -1 \\ 2 \end{pmatrix}$ und $\mathcal{B} = (w^1, w^2, w^3)$, $w^1 = \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix}$, $w^2 = \begin{pmatrix} -1 \\ 2 \\ 0 \end{pmatrix}$, $w^3 = \begin{pmatrix} 0 \\ 2 \\ 1 \end{pmatrix}$.

13 Lineare Gleichungssysteme und affine Unterräume

Es sei K ein Körper. Weiter sei $A \in K^{m,n}$ eine $m \times n$ -Matrix mit den Koeffizienten a_{ij} und $b \in K^m$ ein Spaltenvektor mit den Koeffizienten b_j . x_1, \dots, x_n seien Unbekannte. Das lineare Gleichungssystem

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n &= b_2 \\ &\vdots \\ a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n &= b_m \end{aligned} \tag{45}$$

kann man jetzt in der Form

$$\begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix}$$

oder noch kürzer in der Form

$$Ax = b \tag{46}$$

schreiben, wobei $x = [x_1, \dots, x_n]$. $x \in K^n$ heißt **Lösung** von (46), wenn $Ax = b$ gilt.

$$L = L(A, b) = \{x \in K^n \mid Ax = b\}$$

ist die Lösungsmenge des Gleichungssystems (46). Das Gleichungssystem (46) heißt **inhomogen**, wenn $b \neq 0$. Das Gleichungssystem

$$Ax = 0 \tag{47}$$

heißt das zu (46) gehörende **homogene Gleichungssystem**.

Offensichtlich ist die Lösungsmenge $L(A, 0)$ von (47) gerade der Kern der Abbildung $A : K^n \rightarrow K^m$. Wir schreiben daher auch

$$L(A, 0) = \ker A.$$

Es folgt nun leicht

Satz 13.1

Es sei $A \in K^{m,n}$ eine Matrix vom Rang r und $V = \ker A = \{x \in K^n \mid Ax = 0\}$ die Lösungsmenge des homogenen Gleichungssystem $Ax = 0$. Dann gilt:

- V ist ein $(n - r)$ -dimensionaler Untervektorraum von K^n .
- Ist $b \in K^m$, $b \neq 0$ und ist $x_0 \in K^n$ eine Lösung des inhomogenen Gleichungssystem $Ax = b$, so gilt:

$$L(A, b) = \{x \in K^n \mid Ax = b\} = \{x_0 + v \mid v \in V\} = x_0 + V.$$

c) Für $b \in K^m$ sind äquivalent:

- (i) $Ax = b$ ist lösbar.
- (ii) b ist eine Linearkombination der Spalten von A , also $b \in \langle a^1, \dots, a^n \rangle =$ im A .
- (iii) Der Rang von A stimmt mit dem Rang der erweiterten Matrix (A, b) überein: $\text{rg}(A, b) = \text{rg}A$.

Beweis: Zu a) Nach Satz 9.14 ist

$$\dim V = \text{crg}(A) = n - \text{rg}(A) = n - r.$$

Zu b) Ist $x \in L(A, b)$, so ist $Ax = b$ und somit (wegen $Ax_0 = b$)

$$A(x - x_0) = Ax - Ax_0 = b - b = 0,$$

also $x - x_0 \in V$ und damit $x \in x_0 + V$. Es gilt somit $L(A, b) \subset x_0 + V$. Ist $v \in V$, so ist $x = x_0 + v \in L(A, b)$, denn $Ax = A(x_0 + v) = Ax_0 + Av = Ax_0 = b$. Damit ist auch $x_0 + V \subset L(A, b)$. \square

Zu c): Ist $Ax = b$, so ist $b = \sum x_j a^j \in \langle a^1, \dots, a^n \rangle$. Ist $b \in \langle a^1, \dots, a^n \rangle$, so ist $\text{rg}(A, b) = \text{rg}A$. Ist $\text{rg}(A, b) = \text{rg}A$, so ist b Linearkombination der Spalten von A , also $b = x_1 a^1 + \dots + x_n a^n$, d.h. $x = [x_1, \dots, x_n]$ ist Lösung von $Ax = b$. \square

Das Lösungsverfahren für ein Gleichungssystem

$$Ax = b$$

kennen wir schon:

Mit dem Gaußschen Eliminationsverfahren, d.h. durch elementare Zeilenumformungen der erweiterten Matrix (A, b) , findet man eine invertierbare Matrix $S \in K^{m,m}$, so dass

$$S(A, b) = \left(\begin{array}{c|c} B & c \\ \hline 0 & c' \end{array} \right),$$

wobei $B \in K^{r,n}$ eine Matrix in reduzierter Zeilenstufenform ist, $c \in K^r$, $c' \in K^{n-r}$ und $r = \text{rg}A$. B ist von maximalem Rang. Es gilt:

$$Ax = b \text{ ist lösbar} \Leftrightarrow \text{rg}(A, b) = r \Leftrightarrow \text{rg} \left(\begin{array}{c|c} B & c \\ \hline 0 & c' \end{array} \right) = r \Leftrightarrow c' = 0.$$

Ist $c' \neq 0$, so ist $L(A, b) = \emptyset$. Ist $c' = 0$, so ergibt sich

$$L(A, b) = L(B, c).$$

$L(B, c)$ kann man leicht beschreiben: Es seien $1 \leq j_1 < \dots < j_r \leq n$ und $1 \leq i_1 < \dots < i_{n-r} \leq n$ so gewählt, dass $\{i_1, \dots, i_{n-r}, j_1, \dots, j_r\} = \{1, \dots, n\}$ und

$$B = \left(\begin{array}{cccccc} 0 & \dots & 0 & b_{1j_1} & * & \dots & * \\ 0 & \dots & & & 0 & b_{2j_2} & * & \dots & * \\ \vdots & & & & & & \ddots & & \\ 0 & \dots & & & & & 0 & b_{rj_r} & * & \dots & * \end{array} \right)$$

mit $b_{1j_1} = \dots = b_{rj_r} = 1$. Da eine reduzierte Zeilenstufenform vorliegt ist $b^{j_1} = e^1, \dots, b^{j_r} = e^r$. Die j_ν -te Spalte b^{j_ν} von B ist also der ν -te Einheitsvektor. Also hat B die Blockgestalt

$$B = \begin{pmatrix} 0 & e^1 & * & e^2 & * \dots * & e^r & * \end{pmatrix}.$$

$\uparrow \qquad \qquad \qquad \uparrow \qquad \qquad \qquad \uparrow$
 j_1 -te Spalte j_2 -te Spalte j_r -te Spalte

x_{j_1}, \dots, x_{j_r} sind die Hauptvariablen und $x_{i_1}, \dots, x_{i_{n-r}}$ sind die freien Variablen.

Damit die Bezeichnungen nicht zu kompliziert werden, nehmen wir ohne Einschränkung der Allgemeinheit an, dass $j_1 = 1, \dots, j_r = r$ gilt (eventuell muss man die Variablen x_1, \dots, x_n umnummerieren). Dann sind x_1, \dots, x_r die Hauptvariablen und x_{r+1}, \dots, x_n die freien Variablen. B hat also die Gestalt

$$B = (E_r, B')$$

mit $B' \in K^{r, n-r}$, $B' = (b_{i, r+j})_{\substack{i=1, \dots, r \\ j=1, \dots, n-r}}$. Es gilt somit für $x = \begin{pmatrix} x' \\ x'' \end{pmatrix}$, $x' \in K^r$, $x'' \in K^{n-r}$:

$$\begin{aligned} Ax = b &\iff Bx = c \iff (E_r, B') \begin{pmatrix} x' \\ x'' \end{pmatrix} = c \iff x' + B'x'' \\ &= c \iff x' = c - B'x''. \end{aligned}$$

Damit haben wir die Lösungsmenge $L(A, b)$ als Graf der Abbildung

$$F : K^{n-r} \longrightarrow K^r, \quad F(x'') := c - B'x''$$

beschrieben:

$$L(A, b) = \text{Graph}(F) = \left\{ x = \begin{pmatrix} x' \\ x'' \end{pmatrix} \in K^n \mid x' = F(x'') \right\}.$$

Hat man freie Variablen x'' gewählt, so liegen die Hauptvariablen x' fest: $x' = c - B'x''$. Ist $x = [x_1, \dots, x_n] = \begin{pmatrix} x' \\ x'' \end{pmatrix}$, so bedeutet $x' = F(x'')$:

$$x_i = c_i - \sum_{j=r+1}^n b_{ij} x_j \quad \text{für } i = 1, \dots, r.$$

In Matrixschreibweise ergibt sich auch

$$L(A, b) = \left\{ x = \begin{pmatrix} c \\ 0 \end{pmatrix} + \begin{pmatrix} -B' \\ E_{n-r} \end{pmatrix} x'' \mid x'' \in K^{n-r} \right\}.$$

Für $x'' = 0$ erhält man die spezielle Lösung

$$x_0 = \begin{pmatrix} c \\ 0 \end{pmatrix} = [c_1, \dots, c_r, 0, \dots, 0]$$

von $Ax = b$. Weiter sei $v^j = \begin{pmatrix} -b^{r+j} \\ e^j \end{pmatrix}$ die j -te Spalte der Matrix $\begin{pmatrix} -B' \\ E_{n-r} \end{pmatrix}$. Dann sind v^1, \dots, v^{n-r} linear unabhängige Vektoren in K^n , und v^i ist eine Lösung von $Bx = 0$, denn $B(v^1, \dots, v^{n-r}) = (E_r, B') \begin{pmatrix} -B' \\ E_{n-r} \end{pmatrix} = -B' + B' = 0$ und somit gilt:

$$\begin{aligned} L(A, b) = L(B, c) &= \{x_0 + \lambda_1 v^1 + \dots + \lambda_{n-r} v^{n-r} \mid \lambda_1, \dots, \lambda_{n-r} \in K\} \\ &= x_0 + \langle v^1, \dots, v^{n-r} \rangle. \end{aligned}$$

Beispiel 13.2 Sei $A \in \mathbb{R}^{4,7}$, $b \in \mathbb{R}^4$, $Ax = b$. Kommt man durch das Gaußsche Eliminationsverfahren auf die reduzierte Zeilenstufenform

$$B = \begin{pmatrix} 0 & 1 & 2 & 0 & 3 & 0 & 1 \\ 0 & 0 & 0 & 1 & 2 & 0 & 2 \\ 0 & 0 & 0 & 0 & 0 & 1 & 4 \end{pmatrix}, \quad c = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix},$$

gilt also

$$S(A, b) = \left(\begin{array}{c|c} B & c \\ \hline 0 & c' \end{array} \right),$$

und ist $c' = 0$, so ist $Ax = b$ lösbar und die Lösungsmenge kann man unmittelbar an der Matrix B und der Spalte c ablesen.

Wir numerieren die Variablen um und erhalten

$$\tilde{B} = \begin{pmatrix} 1 & 0 & 0 & | & 0 & 2 & 3 & 1 \\ 0 & 1 & 0 & | & 0 & 0 & 2 & 2 \\ 0 & 0 & 1 & | & 0 & 0 & 0 & 4 \end{pmatrix} = (E_3, B'),$$

$$\begin{pmatrix} -B' \\ E_4 \end{pmatrix} = \begin{pmatrix} 0 & -2 & -3 & -1 \\ 0 & 0 & -2 & -2 \\ 0 & 0 & 0 & -4 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Dann ist

$$\tilde{x}_0 = [1, 2, 3, 0, 0, 0, 0]$$

eine spezielle Lösung von $\tilde{B}x = c$, die vier Spalten

$$\tilde{v}^1 = [0, 0, 0, 1, 0, 0, 0], \dots, \tilde{v}^4 = [-1, -2, -4, 0, 0, 0, 1]$$

von $\begin{pmatrix} -B' \\ E_4 \end{pmatrix}$ sind linear unabhängig, und es gilt:

$$L(\tilde{B}, c) = \tilde{x}_0 + \langle \tilde{v}^1, \tilde{v}^2, \tilde{v}^3, \tilde{v}^4 \rangle.$$

Um die Lösungen von $Ax = b$ zu erhalten, muss man nur die alte Nummerierung wieder herstellen:

$$L(A, b) = L(B, c) = x_0 + \langle v^1, v^2, v^3, v^4 \rangle$$

mit

$$\begin{aligned} x_0 &= [0, 1, 0, 2, 0, 3, 0], \\ v^1 &= [1, 0, 0, 0, 0, 0, 0], \\ v^2 &= [0, -2, 1, 0, 0, 0, 0], \quad \text{usw.} \end{aligned}$$

Auf $\tilde{x}_0, \tilde{v}^1, \dots, \tilde{v}^4$ muss man also die Permutation

$$[x_1, x_2, x_3, x_4, x_5, x_6, x_7] \longmapsto [x_4, x_1, x_5, x_2, x_6, x_3, x_7]$$

anwenden.

Beispiel 13.3 Die Lösungsmenge von

$$Bx = \begin{pmatrix} 1 & 0 & 2 & -2 \\ 0 & 1 & -1 & 3 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

hat die Basis

$$([-2, 1, 1, 0], [2, -3, 0, 1]),$$

das sind die Spalten der Matrix

$$\begin{pmatrix} -2 & 2 \\ 1 & -3 \\ 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} -B' \\ E_2 \end{pmatrix},$$

wobei $B = (E_2, B')$ ist.

Definition 13.4 Eine Teilmenge $L \subset K^n$ heißt **k -dimensionaler affiner Unterraum** $:\Leftrightarrow \exists x_0 \in K^n, V \subset K^n$ k -dimensionaler Untervektorraum, so dass

$$L = x_0 + V$$

gilt. V heißt der zu L gehörige Untervektorraum. $\text{codim} L := n - k$ heißt die **Kodimension** von L in K^n . Hier ist $0 \leq k \leq n$, $k = \dim L$. Die leere Menge \emptyset wird auch als affiner Unterraum angesehen. Ihre Dimension bleibt unbestimmt. Man sagt nur

$$\dim \emptyset < 0 \quad \text{oder} \quad \text{codim} \emptyset > n.$$

Ist $\text{codim} L = 1$, so heißt L eine **Hyperebene** in K^n .

Satz 13.5 Es sei $L \subset K^n$. Dann gilt:

L ist affiner Unterraum von $K^n \iff \exists$ Matrix $A \in K^{m,n}$ und ein $b \in K^m$, so dass $L = L(A, b)$.

Beweis: “ \Leftarrow ” haben wir schon in Satz 13.1 bewiesen.

“ \Rightarrow ” Es sei $V \subset K^n$ ein Untervektorraum und $x_0 \in K^n$, $L = x_0 + V$. Es sei $k = \dim V$. Wir wählen einen zu V komplementären Untervektorraum $U \subset K^n$. Es gilt $\dim U = n - k$. $\pi : K^n \rightarrow U$ sei die Projektion $\pi(v + u) = u$, $v \in V$, $u \in U$. Weiter sei $\alpha : U \rightarrow K^{n-k}$ ein Isomorphismus. Es sei

$$A = \alpha \circ \pi : K^n \rightarrow K^{n-k}$$

A ist eine $(n - k) \times n$ -Matrix und es gilt

$$\ker A = \ker \pi = V.$$

Mit $b := Ax_0$ gilt dann

$$L(A, b) = \{x \in K^n \mid Ax = b\} = \{x \in K^n \mid A(x - x_0) = 0\} = x_0 + V = L.$$

Ist $L = \emptyset$, so wähle man ein nicht lösbares Gleichungssystem. \square

Bemerkung 13.6 Ist $L = x_0 + V$ und (b^1, \dots, b^k) eine Basis von V , $b^j = [b_{1j}, \dots, b_{nj}]$, $j = 1, \dots, k$, $B \in K^{n,k}$ die Matrix aus den Spaltenvektoren b^1, \dots, b^k , so löse man das homogene lineare Gleichungssystem

$$xB = 0.$$

Da $\text{rg} B = k$, gibt es $n - k$ linear unabhängige Lösungen $a_1, \dots, a_{n-k} \in K^{1,n}$. Für

die Matrix $A = \begin{pmatrix} a_1 \\ \vdots \\ a_{n-k} \end{pmatrix} \in K^{n-k,n}$ gilt dann:

$$AB = 0,$$

Dies bedeutet $Ab^j = 0$ für $j = 1, \dots, k$. Da $\text{rg} A = n - k$ gilt, ist

$$V = L(A, 0) = \ker A$$

und

$$L = L(A, b),$$

wobei $b := Ax_0$.

Dies ist ein konstruktives Verfahren zum Auffinden von *Gleichungen* für einen in *Parameterdarstellung* $L = x_0 + \langle b^1, \dots, b^k \rangle$ gegebenen affinen Unterraum von K^n .

Beispiel 13.7

$$L = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} + \left\langle \begin{pmatrix} 2 \\ 1 \\ 2 \end{pmatrix} \right\rangle, \quad B = \begin{pmatrix} 2 \\ 1 \\ 2 \end{pmatrix}, \quad xB = 0 \quad (\text{d.h., } 2x_1 + x_2 + 2x_3 = 0)$$

hat die zwei linear unabhängigen Lösungen

$$a_1 = (1, 0, -1), \quad a_2 = (1, -2, 0).$$

Also ist

$$A = \begin{pmatrix} 1 & 0 & -1 \\ 1 & -2 & 0 \end{pmatrix} \in K^{2,3}$$

eine mögliche Matrix und setzt man

$$b = A \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} = \begin{pmatrix} -2 \\ -3 \end{pmatrix},$$

so ist L die Lösungsmenge von

$$Ax = b, \quad \text{d.h.} \quad \begin{array}{l} x_1 - x_3 = -2 \\ x_1 - 2x_2 = -3 \end{array}.$$

L ist der Durchschnitt der beiden Ebenen

$$E_1 = \{x \in K^3 \mid x_1 - x_3 = -2\} \text{ und } E_2 = \{x \in K^3 \mid x_1 - 2x_2 = -3\}.$$

Satz 13.8 Sind $L_1, L_2 \subset K^n$ affine Unterräume, so ist auch $L_1 \cap L_2$ ein affiner Unterraum, und es gilt: $L_1 \cap L_2 = \emptyset$ oder

$$\text{codim}(L_1 \cap L_2) \leq \text{codim}L_1 + \text{codim}L_2. \quad (48)$$

Beweis: Nach Satz 13.5 kann man Matrizen

$$A_1 \in K^{m_1, n}, \quad A_2 \in K^{m_2, n}, \quad b_1 \in K^{m_1}, \quad b_2 \in K^{m_2}$$

finden, so dass

$$L_i = L(A_i, b_i), \quad \text{codim}L_i = m_i \quad \text{für } i = 1, 2.$$

Dann ist

$$L_1 \cap L_2 = L(A, b),$$

wobei

$$A = \begin{pmatrix} A_1 \\ A_2 \end{pmatrix} \in K^{m_1+m_2, n}, \quad b = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} \in K^{m_1+m_2}.$$

Es gilt also $L_1 \cap L_2 = \emptyset$ oder

$$\text{codim}(L_1 \cap L_2) = \text{rg}A \leq m_1 + m_2 = \text{codim}L_1 + \text{codim}L_2.$$

□

Definition 13.9 Es seien $L_1, L_2 \subset K^n$ affine Unterräume, V_1, V_2 ihre zugehörigen Untervektorräume. Man sagt:

L_1 und L_2 sind in **allgemeiner Lage**, wenn folgendes gilt:

$$\text{codim}(L_1 \cap L_2) = \text{codim}L_1 + \text{codim}L_2,$$

falls $\text{codim}L_1 + \text{codim}L_2 \leq n$ und

$$L_1 \cap L_2 = \emptyset \quad \text{und} \quad V_1 \cap V_2 = 0,$$

falls $\text{codim}L_1 + \text{codim}L_2 > n$.

Bemerkung 13.10

Ist $\text{codim}L_1 + \text{codim}L_2 \leq n$ und sind L_1 und L_2 in allgemeiner Lage, so ist $L_1 \cap L_2 \neq \emptyset$ und $L_1 \cap L_2$ besitzt die kleinstmögliche Dimension, nämlich

$$\dim(L_1 \cap L_2) = \dim L_1 + \dim L_2 - n,$$

weil $\text{codim}(L_1 \cap L_2) = \text{codim}L_1 + \text{codim}L_2$ gilt. Man sagt dann auch: L_1 und L_2 *schneiden* sich in allgemeiner Lage.

Beispiel 13.11

- a) Sind L_1, L_2 3-dimensionale affine Unterräume in \mathbb{R}^5 , so gilt
 $L_1 \cap L_2 = \emptyset$ oder $2 \leq \text{codim}(L_1 \cap L_2) \leq \text{codim}L_1 + \text{codim}L_2 = 4$.
 $\text{codim}(L_1 \cap L_2) = 4$ bedeutet allgemeine Lage.

- b) Ist L_1 3-dimensionaler und L_2 2-dimensionaler affiner Unterraum in \mathbb{R}^5 , so gilt

$$L_1 \cap L_2 = \emptyset \text{ oder } 3 \leq \text{codim}(L_1 \cap L_2) \leq \text{codim } L_1 + \text{codim } L_2 = 5.$$

$$\text{Spezielle Lage: } \begin{cases} 0.) & L_1 \cap L_2 = \emptyset \\ 1.) & L_2 \subset L_1 \\ 2.) & \dim(L_1 \cap L_2) = 1 \end{cases}.$$

Allgemeine Lage: $L_1 \cap L_2 = \{P\}$, P Schnittpunkt.

- c) Es seien L_1, L_2 zwei Geraden in \mathbb{R}^3 . Dann sind L_1, L_2 genau dann in allgemeiner Lage, wenn sie *windschief* sind, d.h. sich nicht schneiden und linear unabhängige Richtungsvektoren haben.

Durch beliebig kleines "Wackeln" wird aus spezieller Lage (d.h. nicht allgemeiner Lage) allgemeine Lage.

Dagegen bleibt die allgemeine Lage bei kleinem Wackeln erhalten.

Das ist anschaulich klar und läßt sich auch exakt beweisen. Dazu muss man auf der Menge der affinen Unterräume von \mathbb{R}^n eine *Topologie* oder einen Abstands begriff einführen, was aber nicht Thema dieser Vorlesung ist.

Desweiteren hat die affine Geometrie einen Schönheitsfehler: Es kann passieren, dass sich zwei affine Teilräume $L_1, L_2 \subset K^n$ nicht schneiden, obwohl $\dim L_1 + \dim L_2 \geq n$ gilt.

Dies kann man beheben, wenn man vom affinen Raum K^n zum projektiven Raum $\mathbb{P}^n(K)$ übergeht, der aus K^n durch Hinzunahme einer sogenannten *unendlichfernen Hyperebene* entsteht. Auch dies soll hier nicht näher untersucht werden.

Übungen

1. Es sei K ein Körper der Charakteristik $p = 13$. Sei

$$L_1 = \langle [1, 2, 3, 0], [1, 1, 2, 1] \rangle \subset K^4$$

und

$$L_2(a) = [0, 1, 2, 12] + \langle [1 - a, 2, 3, a], [1, 1 + a, 2 + a, 1 - a] \rangle.$$

Berechnen Sie $L_1 \cap L_2(a)$. Für welche $a \in K$ schneiden sich L_1 und $L_2(a)$ in allgemeiner Lage?

2. Lösen Sie das Gleichungssystem

$$Ax = b \quad \text{in } \mathbb{C}^4$$

für

$$A = \begin{pmatrix} -1 & 1 - 2i & -1 & 1 - i \\ 1 - i & 2i & 1 & 0 \\ 1 & i & 0 & 1 + i \end{pmatrix}, \quad b = \begin{pmatrix} 3 - i \\ -3 + i \\ 0 \end{pmatrix}.$$

3. Es sei V der von

$$b^1 = [1, 0, 3, -2, 1], \quad b^2 = [2, 4, 1, 0, -2], \quad b^3 = [1, 4, -2, 2, -5]$$

erzeugte Untervektorraum von \mathbb{R}^5 und

$$x_0 = [1, 1, 0, 0, 1], \quad L = x_0 + V.$$

Konstruieren Sie ein lineares Gleichungssystem

$$Ax = b,$$

das L als Lösungsmenge besitzt.

4. Es sei $A \in K^{n,n}$, so dass für alle k mit $1 \leq k \leq n$ die Untermatrix

$$A_{(1,\dots,k)(1,\dots,k)} \in K^{k,k}$$

von A invertierbar ist. Zeigen Sie: Es gibt eine untere Dreiecksmatrix

$$D = \begin{pmatrix} d_1 & & 0 \\ & \ddots & \\ * & & d_n \end{pmatrix} \in K^{n,n}$$

und eine obere Dreiecksmatrix

$$D' = \begin{pmatrix} d'_1 & & * \\ & \ddots & \\ 0 & & d'_n \end{pmatrix} \in K^{n,n},$$

so dass $A = DD'$ (Hinweis: Induktion nach n).

5. Es sei K ein endlicher Körper mit q Elementen. Dann gibt es

$$q^{\binom{n}{2}}(q-1)(q^2-1)\dots(q^n-1)$$

verschiedene invertierbare Matrizen $A \in K^{n,n}$.

6. Zu jedem affinen Unterraum $L \subset K^n$ der Kodimension r gibt es Linearformen $\varphi_1, \dots, \varphi_r \in (K^n)^*$ und Elemente $b_1, \dots, b_r \in K$, so dass

$$L = \bigcap_{i=1}^r \varphi_i^{-1}(\{b_i\})$$

gilt.

14 Ringe, Algebren, Polynomringe, Matrizenalgebren

In diesem Abschnitt werden einige neue algebraische Grundbegriffe eingeführt und an wichtigen Beispielen erläutert.

Definition 14.1 Ein **Ring** (mit Eins) ist eine Menge R mit einer Addition $+$: $R \times R \rightarrow R$ und einer Multiplikation \cdot : $R \times R \rightarrow R$, so dass gilt:

- (1) $(x + y) + z = x + (y + z) \forall x, y, z \in R$.
- (2) $x + y = y + x \forall x, y \in R$.
- (3) $\exists 0 \in R \forall x \in R : x + 0 = x$.
- (4) $\forall x \in R \exists -x \in R : x + (-x) = 0$.
- (5) $(xy)z = x(yz) \forall x, y, z \in R$.
- (6) $(x + y)z = xz + yz$ und $x(y + z) = xy + xz \forall x, y, z \in R$.
- (7) $\exists 1_R \in R$, so dass $1_R x = x 1_R = x \forall x \in R$.

Ein Ring R heißt

- a) **kommutativ**, wenn $xy = yx \forall x, y \in R$.
- b) **nullteilerfrei**, wenn für $x, y \in R$ gilt:
Aus $xy = 0$ folgt $x = 0$ oder $y = 0$.
- c) **Integritätsbereich**, wenn (a), (b) gelten und $R \neq \{0\}$.

Beispiel 14.2

- a) Jeder Körper ist ein Integritätsbereich.
- b) \mathbb{Z} ist ein Integritätsbereich.
- c) Dieses Beispiel ist besonders wichtig: Es sei R ein Ring und $n \in \mathbb{N}$, $n > 0$.

$$M_n(R) = R^{n,n} = \{A = (a_{ij})_{i,j=1,\dots,n} \mid a_{ij} \in R\}$$

ist ein Ring mit der Addition

$$(a_{ij}) + (b_{ij}) := (a_{ij} + b_{ij})$$

und der Multiplikation

$$(a_{ij}) \cdot (b_{ij}) := (c_{ij}), \text{ wobei } c_{ij} := \sum_{k=1}^n a_{ik} b_{kj}.$$

Beweis: (1) - (4) und (6) übertragen sich ohne Mühe von R auf $M_n(R)$.

Zu (5): Sei $A = (a_{ij})$, $B = (b_{ij})$, $C = (c_{ij})$. Es sei $(AB)C = (d_{ij})$. Da $AB = \left(\sum_{k=1}^n a_{ik}b_{kj} \right)_{i,j}$, gilt:

$$\begin{aligned} d_{ij} &= \sum_{l=1}^n \left(\sum_{k=1}^n a_{ik}b_{kl} \right) c_{lj} \stackrel{(6)}{=} \sum_{l=1}^n \sum_{k=1}^n (a_{ik}b_{kl})c_{lj} \\ &\stackrel{(5)}{=} \sum_{l=1}^n \sum_{k=1}^n a_{ik}(b_{kl}c_{lj}) \stackrel{(2)}{=} \sum_{k=1}^n \sum_{l=1}^n a_{ik}(b_{kl}c_{lj}) \\ &\stackrel{(6)}{=} \sum_{k=1}^n a_{ik} \sum_{l=1}^n b_{kl}c_{lj}; \end{aligned}$$

und somit $(AB)C = A(BC)$.

Zu (7): Die Einheitsmatrix E_n ist das Einselement. \square

$M_n(R)$ heißt der **Matrizenring n -ten Grades** über dem Ring R .

- d) Ist R ein Ring, X eine Menge, so ist auch Abb $(X, R) = \{f|f : X \rightarrow R\}$ ein Ring mit den Verknüpfungen

$$\begin{aligned} (f+g)(x) &:= f(x) + g(x) \\ (fg)(x) &:= f(x)g(x). \end{aligned}$$

Definition 14.3 Ist S ein Ring und $R \subset S$, so heißt R **Unterring** von S , wenn gilt:

- $x, y \in R \Rightarrow x - y \in R$.
- $x, y \in R \Rightarrow xy \in R$.
- $1 \in R$.

R ist dann ebenfalls ein Ring mit den von S auf R eingeschränkten Verknüpfungen (Übung).

Beispiel 14.4

- $\mathbb{Z} \subset \mathbb{Q}$ ist Unterring.
- Sei R ein kommutativer Ring mit Eins, S ein Ring, so dass $R \subset S$ Unterring ist. Sei $a \in S$, und es gelte $ar = ra \forall r \in R$. Dann gilt:
 - $R[a] = \{b \in S \mid \exists n \in \mathbb{N}, a_0, \dots, a_n \in R, \text{ so dass } b = a_0 + a_1a + a_2a^2 + \dots + a_na^n\}$ ist ein Unterring von S . $R[a]$ heißt der von R und a erzeugte Unterring von S .
 - $R[a]$ ist kommutativ und enthält R als Unterring.

Beweis: Zu (i): Seien $b, c \in R[a]$, $b = \sum_{\nu=0}^n a_\nu a^\nu$, $c = \sum_{\nu=0}^n a'_\nu a^\nu$, dann ist

$$b - c = \sum_{\nu=0}^n (a_\nu - a'_\nu) a^\nu \in R[a] \text{ und}$$

$$\begin{aligned} bc &= \left(\sum_{\nu=0}^n a_\nu a^\nu \right) \left(\sum_{\mu=0}^n a'_\mu a^\mu \right) \\ &= \sum_{\nu,\mu=0}^n a_\nu a^\nu a'_\mu a^\mu = \sum_{\nu,\mu=0}^n a_\nu a'_\mu a^{\nu+\mu} \in R[a]. \end{aligned}$$

$$\text{Zu (ii): } bc = \sum_{\nu,\mu=0}^n a_\nu a'_\mu a^{\nu+\mu} = \sum_{\nu,\mu=0}^n a'_\mu a_\nu a^{\nu+\mu} = cb. \quad \square$$

Konkrete Beispiele sind $\mathbb{Z}[i] = \{a + bi \mid a, b \in \mathbb{Z}\} \subset \mathbb{C}$, der Ring der ganzen Gaußschen Zahlen, und das schon bekannte Beispiel $\mathbb{Q}[\sqrt{d}] \subset \mathbb{C}$, wobei $d \in \mathbb{Z}$ mit $\sqrt{d} \notin \mathbb{Q}$.

Definition 14.5 Es sei K ein Körper und R eine Menge mit drei Verknüpfungen, einer Addition $+$: $R \times R \rightarrow R$, einer skalaren Multiplikation \cdot : $K \times R \rightarrow R$ und einer Multiplikation \circ : $R \times R \rightarrow R$. Dann heißt $(R, +, \circ, \cdot)$ eine **K -Algebra (mit Eins)**, falls gilt:

- $(R, +, \cdot)$ ist ein K -Vektorraum.
- $(R, +, \circ)$ ist ein Ring mit Einselement 1_R .
- Für alle $A, B \in R$, $a \in K$ gilt: $a(A \circ B) = (aA) \circ B = A \circ (aB)$.

Als wichtigste Beispiele haben wir

Satz und Definition 14.6 Es sei V ein K -Vektorraum. Eine lineare Abbildung $\varphi : V \rightarrow V$ heißt **Endomorphismus** von V . Die Menge

$$\text{End } V = \text{End}_K V := \text{Hom}(V, V)$$

ist mit der Komposition von Endomorphismen als Multiplikation eine K -Algebra mit Einselement id_V . $\text{End } V$ heißt die **Endomorphismenalgebra** (auch: der **Endomorphismenring**) von V .

Ein weiteres Beispiel ist der Matrizenring $M_n(K)$ n -ten Grades über einem Körper K . Die skalare Multiplikation ist mit der Matrizenmultiplikation verträglich:

$$a(AB) = (aA)B = A(aB) \quad \text{für alle } a \in K, A, B \in M_n(K).$$

$M_n(K)$ heißt deshalb auch die (**volle**) **Matrizenalgebra n -ten Grades** über dem Körper K .

Definition 14.7 Sei S eine K -Algebra und $R \subset S$. R heißt **K -Unteralgebra** von S , falls gilt:

- R ist K -Untervektorraum.

b) R ist Unterring von S .

Beispiel 14.8 $M_2(K)$ ist 4-dimensionale K -Algebra.

a) $M_2(K)$ ist nicht kommutativ; z.B. ist

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

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

und $\begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix} \neq \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}$ für jeden Körper K .

b) Sei

$$D = \begin{pmatrix} 0 & 1 \\ d & 0 \end{pmatrix}.$$

Dann gilt $D^2 = dE$, wobei $E = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$.

$K[D] = \{aE + bD \mid a, b \in K\}$ ist eine zweidimensionale K -Unteralgebra von $M_2(K)$. Es gibt drei Fälle.

1. Fall: $d \neq a^2 \forall a \in K$, d.h. es gibt keine Quadratwurzel von d in K . Dann ist für $(a, b) \neq (0, 0)$

$$(aE + bD)(aE - bD) = (a^2 - b^2d)E \neq 0.$$

Hieraus folgt sofort: $K[D]$ ist ein Erweiterungskörper von K , wobei K mit

$$KE = \left\{ \begin{pmatrix} a & 0 \\ 0 & a \end{pmatrix} \mid a \in K \right\}$$

identifiziert wird. Als konkretes Beispiel wählen wir $K = \mathbb{F}_3$ und $d = [2]$. Dann ist d kein Quadrat in \mathbb{F}_3 , weil $[0]^2 = [0]$, $[1]^2 = [1] = [2]^2$. Wir bezeichnen die Elemente von \mathbb{F}_3 einfach mit $0, 1, 2$.

$$L = \mathbb{F}_3 \left[\begin{pmatrix} 0 & 1 \\ 2 & 0 \end{pmatrix} \right] = \left\{ \begin{pmatrix} a & b \\ 2b & a \end{pmatrix} \mid a, b \in \mathbb{F}_3 \right\}$$

ist also ein Körper mit neun Elementen. Übung: Es gibt eine Matrix $A \in L$, so dass alle von Null verschiedenen Elemente in L Potenzen von A sind. Lösung:

Zum Beispiel $\begin{pmatrix} 1 & 1 \\ 2 & 1 \end{pmatrix}$.

2. Fall: $d = q^2$ für ein $q \in K \setminus \{0\}$. Es folgt:

$$(qE + D)(qE - D) = q^2E - D^2 = (q^2 - d)E = 0;$$

und somit ist $qE + D$ ein **Nullteiler** von $K[D]$.

3. Fall: $d = 0$. Sei $N := D = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$. Es gilt: $N^2 = 0$. N ist **nilpotent**. $K[N] = \{aE + bN \mid a, b \in K\}$ besteht ausschließlich aus Elementen die entweder invertierbar oder nilpotent sind.

Behauptung: $S = aE + bN$ ist genau dann invertierbar, wenn $a \neq 0$ ist.

Beweis: Sei $a \neq 0$. Dann ist $S = a \left(E + \frac{b}{a}N \right)$; und somit $S^{-1} = a^{-1} \left(E - \frac{b}{a}N \right)$, weil

$$SS^{-1} = E - \left(\frac{b}{a} \right)^2 N^2 = E.$$

Ist $a = 0$, so ist $S = bN$ nilpotent.

Definition 14.9

a) Es seien R, S Ringe (mit Eins). Eine Abbildung $\varphi : R \rightarrow S$ heißt **Ringhomomorphismus**, wenn gilt:

(i) $\varphi(x + y) = \varphi(x) + \varphi(y) \quad \forall x, y \in R.$

(ii) $\varphi(xy) = \varphi(x)\varphi(y) \quad \forall x, y \in R.$

(iii) $\varphi(1_R) = 1_S.$

Ein bijektiver Ringhomomorphismus heißt **Ringisomorphismus**.

b) Es sei K ein Körper; R, S seien K -Algebren. Eine Abbildung $\varphi : R \rightarrow S$ heißt **K -Algebrahomomorphismus**, wenn φ ein K -linearer Ringhomomorphismus ist. Ein bijektiver K -Algebrahomomorphismus heißt **K -Algebraisomorphismus**.

Beispiel 14.10

a) Ist R ein Ring, so gibt es genau einen Ringhomomorphismus $\varphi : \mathbb{Z} \rightarrow R$, nämlich $\varphi(n) = n1_R$.

b) Ist R eine K -Algebra, so gibt es genau einen K -Algebrahomomorphismus $\varphi : K \rightarrow R$, nämlich $\varphi(a) = a1_R$.

Für $R = M_n(K)$ ist dies die Abbildung

$$a \mapsto \begin{pmatrix} a & & 0 \\ & \ddots & \\ 0 & & a \end{pmatrix} = aE_n.$$

Für $R = \text{Abb}(X, K)$ ist dies die Abbildung

$$a \mapsto (\text{konstante Funktion } x \mapsto a).$$

$\varphi : K \rightarrow R$ heißt die **kanonische Abbildung**. Ist $R \neq 0$, d.h. gilt $0 \neq 1_R$, so ist die Abbildung $\varphi : K \rightarrow R$, $\varphi(a) = a1_R$ injektiv, denn wäre $\ker \varphi \neq 0$, so gäbe es ein $a \in K$, $a \neq 0$ mit $a1_R = 0$; dann wäre auch

$$0 = a^{-1}(a1_R) = (a^{-1}a)1_R = 1 \cdot 1_R = 1_R.$$

Jede von Null verschiedene K -Algebra R enthält die zu K isomorphe Unter-
algebra $K1_R = \{a1_R | a \in K\}$, und es gilt nach Definition 14.5 c):

$$(a1_R)A = aA = A(a1_R)$$

für alle $A \in R$ und für alle $a \in K$. Nach Beispiel 14.4 b) ist daher für jedes
 $A \in R$ die Menge

$$K[A] := K1_R[A] = \left\{ \sum_{\nu=0}^n a_\nu A^\nu \mid n \in \mathbb{N}, a_0, \dots, a_n \in K \right\}$$

ein Unterring von R und natürlich auch ein K -Untervektorraum, also eine
 K -Unteralgebra von R . Außerdem ist $K[A]$ kommutativ.

Ist V ein 2-dimensionaler K -Vektorraum und $\varphi \in \text{End } V$, so gibt es Elemente
 $a, b \in K$, so dass

$$\varphi^2 + a\varphi + b \text{id}_V = 0$$

gilt. Jede Potenz φ^ν , $\nu \geq 2$, kann man also als Linearkombination von φ und
 id_V schreiben. Ist $\varphi \neq a \text{id}_V$ für alle $a \in K$, so ist

$$K[\varphi] \subset \text{End } V$$

eine 2-dimensionale K -Unteralgebra von $\text{End } V$.

- c) Es sei K ein Körper und R die K -Algebra aller Funktionen $f : K \rightarrow K$. Mit
 f_1 bezeichnen wir die Funktion $f_1(x) = x$. Dann ist

$$K[f_1] \subset R$$

die K -Unteralgebra der Polynomfunktionen auf K .

Wir geben nun die abstrakte Definition des Polynomrings.

Definition 14.11 Es sei R ein Ring und X eine Unbestimmte (d.h. X ist ein for-
males Symbol).

Ein **Polynom** in der Unbestimmten X mit Koeffizienten in R ist ein formaler Aus-
druck

$$P = a_0 + a_1X + a_2X^2 + \dots + a_nX^n = \sum_{\nu=0}^n a_\nu X^\nu,$$

wobei $a_0, \dots, a_n \in R$, $n \in \mathbb{N}$ und $X^0 := 1_R$, $X^1 := X$.

Ist $Q = b_0 + b_1X + \dots + b_mX^m$ ein weiteres Polynom, so wird definiert

$$\begin{aligned} P = Q \iff \forall 0 \leq i \leq \min(n, m) : & a_i = b_i \quad \text{und} \\ \forall \min(n, m) < i \leq n : & a_i = 0 \quad \text{und} \\ \forall \min(n, m) < i \leq m : & b_i = 0. \end{aligned}$$

Mit $R[X]$ bezeichnen wir die Menge aller Polynome in X mit Koeffizienten in R .
Weiter sei $R^{(\mathbb{N})}$ die Menge aller Folgen

$$a = (a_k)_{k \in \mathbb{N}} = (a_0, a_1, a_2, a_3, \dots)$$

von Elementen $a_k \in R$ mit der Eigenschaft:

$$\exists n \in \mathbb{N}, \text{ so dass } a_k = 0 \text{ f\u00fcr alle } k > n.$$

Nach Definition der Gleichheit von Polynomen ist die Abbildung

$$a = (a_0, a_1, a_2, \dots) \mapsto P = a_0 + a_1X + a_2X^2 + \dots$$

eine Bijektion $R^{(\mathbb{N})} \rightarrow R[X]$.

a hei\u00dft die **Koeffizientenfolge** von P .

Die Gleichheit von Polynomen pr\u00fcft man durch Koeffizientenvergleich.

Die **Addition** und **Multiplikation** von Polynomen wird so definiert: Seien $P, Q \in R[X]$, $a = (a_0, a_1, a_2, \dots)$ sei die Koeffizientenfolge von P und $b = (b_0, b_1, b_2, \dots)$ sei die Koeffizientenfolge von Q . Dann ist $P + Q$ das Polynom mit

$$a + b := (a_0 + b_0, a_1 + b_1, a_2 + b_2, \dots)$$

als Koeffizientenfolge, und PQ ist das Polynom mit der Koeffizientenfolge

$$a * b := \left(a_0b_0, a_0b_1 + a_1b_0, \dots, \sum_{\nu+\mu=k} a_\nu b_\mu, \dots \right).$$

Es gilt also nach Definition:

$$P + Q = (a_0 + b_0) + (a_1 + b_1)X + (a_2 + b_2)X^2 + \dots = \sum_{\nu=0}^n (a_\nu + b_\nu)X^\nu$$

(falls $a_i = b_i = 0$ f\u00fcr $i > n$) und

$$PQ = a_0b_0 + (a_0b_1 + a_1b_0)X + \dots + \left(\sum_{\nu+\mu=k} a_\nu b_\mu \right) X^k + \dots + a_n b_m X^{n+m}$$

(falls $a_i = 0$ f\u00fcr $i > n$, $a_n \neq 0$, $b_j = 0$ f\u00fcr $j > m$, $b_m \neq 0$).

Satz 14.12

- $R[X]$ ist ein Ring, und $R \subset R[X]$ ist Unterring.
- Ist R kommutativ, so ist auch $R[X]$ kommutativ.
- Ist R nullteilerfrei, so ist $R[X]$ nullteilerfrei.
- Ist K ein K\u00f6rper, so ist $K[X]$ eine unendlich-dimensionale kommutative K -Algebra. $(1, X, X^2, X^3, \dots)$ ist eine K -Vektorraumbasis von $K[X]$.

Beweis zu a): Wir m\u00fcssen die Axiome (1) - (7) aus Definition 14.1 nachpr\u00fcfen.

(1) und (2) sind klar.

Zu (3): Das Nullpolynom $0 \in K[X]$ ist das Polynom mit der Folge $(0, 0, \dots)$ als Koeffizientenfolge. Dann gilt offensichtlich:

$$P + 0 = P \text{ f\u00fcr alle } P \in K[X].$$

Zu (4): Ist $P = a_0 + a_1X + a_2X^2 + \dots + a_nX^n$, so definiert man

$$-P := (-a_0) + (-a_1)X + \dots + (-a_n)X^n.$$

Dann gilt:

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

Zu (5): Es sei $P = \sum a_\nu X^\nu$, $Q = \sum b_\nu X^\nu$, $R = \sum c_\nu X^\nu$. Dann gilt:

$$\begin{aligned} P(QR) &= \left(\sum_\nu a_\nu X^\nu \right) \left(\sum_k \left(\sum_{\nu+\mu=k} b_\nu c_\mu \right) X^k \right) \\ &= \sum_m \left(\sum_{l+k=m} a_l \sum_{\nu+\mu=k} b_\nu c_\mu \right) X^m \\ &= \sum_m \left(\sum_{l+\nu+\mu=m} a_l b_\nu c_\mu \right) X^m. \end{aligned}$$

Berechnet man $(PQ)R$, so erhält man dasselbe Ergebnis.

Zu (6): P, Q, R seien wie im Beweis zu (5) gewählt. Dann gilt:

$$\begin{aligned} P(Q+R) &= \sum_k \left(\sum_{\nu+\mu=k} a_\nu (b_\mu + c_\mu) \right) X^k \\ &= \sum_k \left(\sum_{\nu+\mu=k} a_\nu b_\mu \right) X^k + \sum_k \left(\sum_{\nu+\mu=k} a_\nu c_\mu \right) X^k \\ &= PQ + PR. \end{aligned}$$

Genauso folgt $(P+Q)R = PR + QR$.

Zu (7): Nach Konstruktion ist $R \subset R[X]$, und das Einselement 1 in R ist auch das Einselement in $R[X]$. Damit ist $R[X]$ ein Ring, und offensichtlich ist R ein Unterring von $R[X]$. \square

Die Elemente aus R heißen auch **konstante Polynome**.

$a \in R$ besitzt als Polynom in $R[X]$ die Koeffizientenfolge $(a, 0, 0, \dots)$.

Beweis zu b): Klar! \square

Beweis zu c): Seien $P, Q \in R[X]$, $P, Q \neq 0$,

$$\begin{aligned} P &= a_0 + a_1X + \dots + a_nX^n \quad \text{mit } a_n \neq 0, \\ Q &= b_0 + b_1X + \dots + b_mX^m \quad \text{mit } b_m \neq 0. \end{aligned}$$

Da R nullteilerfrei ist, ist $a_nb_m \neq 0$, also auch $PQ = a_0b_0 + \dots + a_nb_mX^{n+m} \neq 0$. \square

Beweis zu d): Ist K ein Körper, so ist $K[X]$ auch ein K -Vektorraum. Die skalare Multiplikation $K \times K[X] \rightarrow K[X]$ entsteht durch Einschränkung der Polynommultiplikation $K[X] \times K[X] \rightarrow K[X]$ auf $K \times K[X]$. Für $a \in K$, $P = \sum_{\nu=0}^n a_\nu X^\nu \in K[X]$

ist

$$aP = \sum_{\nu=0}^n (aa_{\nu})X^{\nu}.$$

Für jedes $n \in \mathbb{N}$ sind X^0, X^1, \dots, X^n nach Konstruktion linear unabhängig über K , und $\{X^0, X^1, X^2, \dots\}$ ist ein Erzeugendensystem von $K[X]$ als K -Vektorraum. \square
 X^n hat die Koeffizientenfolge

$$(0, 0, \dots, 0, 1, 0, 0, \dots) = (\delta_{n0}, \delta_{n1}, \dots, \delta_{nn}, \delta_{n,n+1}, \dots),$$

und nach Definition der Multiplikation gilt:

$$X^{n+1} = X X^n.$$

Das Polynom (genauer: das **Monom**) X^n ist somit die n -te Potenz von X .

Definition 14.13 Ist R ein Ring, $P = \sum_{\nu=0}^n a_{\nu}X^{\nu} \in R[X]$, $P \neq 0$ und $a_n \neq 0$, so heißt $\text{grad } P := n$ der **Grad** von P und a_n der **Leitkoeffizient** von P . Dagegen heißt a_0 der **konstante Term** von P (bzgl. X). Für das Nullpolynom wird $\text{grad } 0 = -\infty$ gesetzt.

Lemma 14.14 Ist R ein Integritätsbereich, so gilt für alle $P, Q \in R[X]$:

- a) $\text{grad } (P + Q) \leq \max(\text{grad } P, \text{grad } Q)$.
- b) $\text{grad } (PQ) = \text{grad } P + \text{grad } Q$.

(Hier wird $a + (-\infty) = (-\infty) + a = -\infty$ für alle $a \in \mathbb{N} \cup \{-\infty\}$ und $-\infty \leq a$ für alle $a \in \mathbb{N} \cup \{-\infty\}$ gesetzt.)

Wichtig ist die folgende fundamentale Eigenschaft des Polynomrings $R[X]$ eines kommutativen Rings R .

Satz 14.15 (*universelle Eigenschaft des Polynomrings*) Es seien R, S Ringe, $\varphi : R \rightarrow S$ sei ein Ringhomomorphismus, und es sei $b \in S$. R sei kommutativ, und es gelte $\varphi(a)b = b\varphi(a)$ für alle $a \in R$. Dann gibt es genau einen Ringhomomorphismus $\Phi : R[X] \rightarrow S$ mit $\Phi(a) = \varphi(a)$ für alle $a \in R$ und $\Phi(X) = b$.

Beweis:

- a) Eindeutigkeit:

Sei $\Phi : R[X] \rightarrow S$ Ringhomomorphismus mit $\Phi(a) = \varphi(a)$ für $a \in R$ und $\Phi(X) = b$. Dann ist

$$\Phi(X^n) = \Phi(XX^{n-1}) = \Phi(X)\Phi(X^{n-1}).$$

Induktiv ergibt sich also $\Phi(X^n) = b^n$ und somit

$$\Phi \left(\sum_{\nu=0}^n a_{\nu}X^{\nu} \right) = \sum_{\nu=0}^n \Phi(a_{\nu})\Phi(X^{\nu}) = \sum_{\nu=0}^n \varphi(a_{\nu})b^{\nu}.$$

Also ist Φ eindeutig bestimmt, und es muss gelten:

$$\Phi \left(\sum_{\nu=0}^n a_{\nu} X^{\nu} \right) = \sum_{\nu=0}^n \varphi(a_{\nu}) b^{\nu}. \quad (49)$$

b) Existenz:

Durch die Formel (49) ist eine Abbildung $\Phi : R[X] \rightarrow S$ definiert. Wir müssen nur noch zeigen, dass Φ ein Ringhomomorphismus ist. Es sei

$$P = \sum_{\nu=0}^n a_{\nu} X^{\nu}, \quad Q = \sum_{\nu=0}^n b_{\nu} X^{\nu}.$$

Dann ist

$$P + Q = \sum_{\nu=0}^n (a_{\nu} + b_{\nu}) X^{\nu};$$

und somit

$$\begin{aligned} \Phi(P + Q) &= \sum_{\nu=0}^n \varphi(a_{\nu} + b_{\nu}) b^{\nu} = \sum_{\nu=0}^n (\varphi(a_{\nu}) + \varphi(b_{\nu})) b^{\nu} \\ &= \sum_{\nu=0}^n \varphi(a_{\nu}) b^{\nu} + \sum_{\nu=0}^n \varphi(b_{\nu}) b^{\nu} = \Phi(P) + \Phi(Q). \end{aligned}$$

Weiter ist

$$PQ = \sum_{k=0}^{2n} \left(\sum_{\nu+\mu=k} a_{\nu} b_{\mu} \right) X^k,$$

also

$$\begin{aligned} \Phi(PQ) &= \sum_{k=0}^{2n} \varphi \left(\sum_{\nu+\mu=k} a_{\nu} b_{\mu} \right) b^k = \sum_{k=0}^{2n} \sum_{\nu+\mu=k} \varphi(a_{\nu}) \varphi(b_{\mu}) b^k \\ &= \left(\sum_{\nu=0}^n \varphi(a_{\nu}) b^{\nu} \right) \left(\sum_{\mu=0}^n \varphi(b_{\mu}) b^{\mu} \right) = \Phi(P) \Phi(Q). \end{aligned}$$

Schließlich ist $\Phi(1) = \varphi(1) = 1$.

□

Als Folgerung erhalten wir

Satz 14.16 Es sei K ein Körper und R eine K -Algebra. Zu jedem $A \in R$ gibt es genau einen K -Algebrahomomorphismus

$$\Phi_A : K[X] \longrightarrow R \quad \text{mit} \quad \Phi_A(X) = A.$$

Beweis: $\varphi : K \rightarrow R$ sei die natürliche Abbildung $\varphi(a) = a1_R$. Dann ist

$$\varphi(a)A = aA = A(a1_R) = A\varphi(a)$$

für alle $a \in K$. Also kann man Satz 14.15 anwenden. Man erhält einen Ringhomomorphismus

$$\Phi_A = \Phi : K[X] \rightarrow R$$

mit $\Phi(a) = a1_R$ und $\Phi(X) = A$; die Bedingung $\Phi(a) = a1_R$ für $a \in K$ bedeutet genau, dass Φ ein K -Algebrahomomorphismus ist. \square

Beispiel 14.17 Es sei K ein Körper, $n \in \mathbb{N}$, $n > 0$.

- a) Für jede Matrix $A \in M_n(K)$ gibt es nach Satz 14.16 einen K -Algebrahomomorphismus

$$\Phi_A : K[X] \rightarrow M_n(K) \quad \text{mit} \quad \Phi_A(X) = A.$$

Dann ist offensichtlich $K[A] = \text{im } \Phi_A$ (vgl. Beispiel 14.10 b).

Im Fall $n = 2$ haben wir schon gesehen, dass die Abbildung $\Phi_A : K[X] \rightarrow K[A]$ nicht injektiv ist. Vielmehr gibt es Elemente $a, b \in K$, so dass

$$A^2 + aA + bE_2 = 0$$

ist. $P = X^2 + aX + b \in K[X]$ ist somit ein Polynom mit

$$\Phi_A(P) = A^2 + aA + bE_2 = 0.$$

Auch für beliebiges n kann Φ_A aus Dimensionsgründen nicht injektiv sein.

- b) Ist $a \in K$, so ist der K -Algebrahomomorphismus

$$\Phi_a : K[X] \rightarrow K \quad \text{mit} \quad \Phi_a(X) = a$$

die **Auswertungsabbildung**

$$P = \sum_{\nu=0}^n a_\nu X^\nu \longmapsto \sum_{\nu=0}^n a_\nu a^\nu \in K.$$

Für X wird a eingesetzt.

Notation 14.18 Statt $\Phi_A(P)$ schreibt man auch $P(A)$. Die Unbestimmte X wird durch A ersetzt.

Die Ringe \mathbb{Z} und $K[X]$ (K -Körper) haben eine sehr verwandte algebraische Struktur. Der Grund liegt in folgendem

Satz 14.19 (Division mit Rest): Es sei K ein Körper, $G \in K[X]$, $G \neq 0$. Dann gilt:

$\forall F \in K[X] \exists_1 Q, R \in K[X]$, so dass $F = QG + R$ und $\text{grad } R < \text{grad } G$.

R heißt der **Rest** von F bei Division durch G und wird mit $F \bmod G$ bezeichnet. Für $F_1, F_2 \in K[X]$ definiert man

$$F_1 \equiv F_2 \bmod G \quad (F_1 \text{ und } F_2 \text{ sind kongruent modulo } G) :\Leftrightarrow \\ F_1 \bmod G = F_2 \bmod G.$$

Es sei $\text{grad } G = d$. Dann ist $F \mapsto F \bmod G$ eine surjektive K -lineare Abbildung

$$K[X] \longrightarrow K[X]_{d-1} = \{\text{Polynome vom Grad } < d\}.$$

Beweis:

1. Existenz:

Sei

$$\begin{aligned} F &= a_0 + a_1X + \dots + a_nX^n, & a_n &\neq 0, \\ G &= b_0 + b_1X + \dots + b_mX^m, & b_m &\neq 0. \end{aligned}$$

Induktion nach n :

- (a) Ist $n = 0$, so setze $R = F$, $Q = 0$, falls $m > 0$, und $R = 0$, $Q = a_0b_0^{-1}$, falls $m = 0$.
- (b) Jetzt sei $n > 0$ und die Existenz für Polynome vom Grad $< n$ schon bewiesen. Ist $n < m$, so können wir $Q = 0$ und $R = F$ setzen. Sei also $m \leq n$. Dann schreiben wir

$$F = a_nb_m^{-1}X^{n-m}G + \tilde{F},$$

und erhalten $\text{grad } \tilde{F} < n$. Nach Induktionsvoraussetzung ist $\tilde{F} = \tilde{Q}G + R$ mit $\text{grad } R < m$, also

$$F = (a_nb_m^{-1}X^{n-m} + \tilde{Q})G + R.$$

2. Eindeutigkeit:

Es gelte: $QG + R = Q'G + R'$ mit $\text{grad } R < \text{grad } G$ und $\text{grad } R' < \text{grad } G$. Dann folgt

$$(Q - Q')G = R' - R,$$

also

$$\text{grad } (Q - Q') + \text{grad } G = \text{grad } (R - R');$$

das geht nur, wenn $Q = Q'$ und $R = R'$ (Beachte: $-\infty + a = -\infty$).

□

Definition 14.20 Es sei $K \subset L$ eine Körpererweiterung und $P \in K[X]$. $\alpha \in L$ heißt **Nullstelle** von P , wenn $P(\alpha) = 0$ gilt.

$\alpha \in L$ heißt **algebraisch** über K , wenn es ein Polynom $P \in K[X]$ mit $\text{grad } P \geq 1$ gibt, so dass

$$P(\alpha) = 0.$$

Beispiel 14.21 $X^2 + 1 \in \mathbb{R}[X]$ hat keine Nullstellen in \mathbb{R} , aber in \mathbb{C} : i und $-i$ sind Nullstellen von $X^2 + 1$. Es gilt: $\mathbb{C} = \mathbb{R}[i]$.

Die Abbildung $\Phi_i : \mathbb{R}[X] \rightarrow \mathbb{C}$ ist surjektiv und

$$\ker \Phi_i = \{F \in \mathbb{R}[X] \mid F \bmod (X^2 + 1) = 0\}.$$

Lemma 14.22 Ist $P \in K[X]$ und $a \in K$ Nullstelle von P , so gilt:

$$P = Q(X - a) \text{ mit } Q \in K[X], \quad \text{grad } Q = \text{grad } P - 1.$$

Ein Polynom $P \in K[X]$ vom Grad n besitzt höchstens n Nullstellen in K .

Lemma 14.23 Es sei K ein Körper und V die kommutative K -Algebra der Polynomfunktionen $f : K \rightarrow K$. Dann ist die kanonische Abbildung

$$\Phi : K[X] \longrightarrow V, \quad P \longmapsto P(f_1) = f$$

(also $f(x) = (\Phi_{f_1}(P))(x) = P(f_1)(x) = P(x)$ für alle $x \in K$) surjektiv.

Aber Φ ist injektiv nur dann, wenn K unendlich ist.

Beweis:

$$\ker \Phi = \{P \in K[X] \mid P(a) = 0 \forall a \in K\}.$$

Also ist $\ker \Phi = 0$, wenn K unendlich ist.

Ist K endlich, so ist

$$P := \prod_{a \in K} (X - a) \in \ker \Phi, \quad \text{aber } P \neq 0.$$

□

Mehr über Polynome folgt später.

Übungen

1. Es sei K ein Körper. Welche der folgenden Teilmengen ist eine K -Unteralgebra der Matrizenalgebra $M_n(K)$?

(a) $R_1 = \{A = (a_{ij}) \mid a_{ij} = 0 \text{ für } i \neq j\}$.

(b) $R_2 = \{A = (a_{ij}) \mid a_{ij} = 0 \text{ für } i < j\}$.

(c) $R_3 = \left\{ A = \begin{pmatrix} A_1 & C \\ 0 & A_2 \end{pmatrix} \mid A_1 \in M_{n_1}(K), A_2 \in M_{n_2}(K), C \in K^{n_1, n_2} \right\},$
wobei $n = n_1 + n_2$.

Bestimmen Sie für $R = M_n(K), R_1, R_2$ die Menge

$$Z(R) = \{A \in R \mid AB = BA \text{ für alle } B \in R\}.$$

2. Es sei R ein Ring. $x \in R$ heißt **nilpotent**, wenn $x^m = 0$ für ein $m \in \mathbb{N}$ gilt. Zeigen Sie:

- (a) Sind $x, y \in R$ vertauschbar (d.h. es gilt $xy = yx$), so ist $x + y$ nilpotent, wenn x und y es sind.

- (b) Finden Sie einen nicht kommutativen Ring R , zwei nilpotente nicht vertauschbare Elemente $x, y \in R$, so dass $x + y$ sogar invertierbar ist, d.h.

$$(x + y)z = z(x + y) = 1_R$$

für ein $z \in R$ gilt.

3. Es sei $A \in M_2(\mathbb{C})$. Beweisen Sie: Es gibt eine invertierbare Matrix $S \in M_2(\mathbb{C})$, so dass gilt:

$$SAS^{-1} = \begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix} \quad \text{mit } a, b \in \mathbb{C}$$

oder

$$SAS^{-1} = \begin{pmatrix} a & 1 \\ 0 & a \end{pmatrix} \quad \text{mit } a \in \mathbb{C}.$$

Gilt die analoge Aussage auch für den Körper \mathbb{R} an Stelle von \mathbb{C} ? Betrachten Sie z.B. $A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$.

4. Es seien $F = X^5 - 4X - 1$, $G = 2X^3 - 2X^2 + 1$, $H = X^3 + X^2 - X + 2 \in \mathbb{F}_5[X]$ gegeben. Teilen Sie F mit Rest durch G und durch H .

15 Determinanten

Beispiel 15.1 Wir haben schon gesehen, dass ein lineares Gleichungssystem

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} e \\ f \end{pmatrix}$$

für $a, b, c, d, e, f \in K$ genau dann eindeutig lösbar ist, wenn $\begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc \neq 0$ ist.

Es gilt:

$$x = \frac{\begin{vmatrix} e & b \\ f & d \end{vmatrix}}{\begin{vmatrix} a & b \\ c & d \end{vmatrix}} \quad \text{und} \quad y = \frac{\begin{vmatrix} a & e \\ b & f \end{vmatrix}}{\begin{vmatrix} a & b \\ c & d \end{vmatrix}}.$$

Analog kann man auch bei drei Gleichungen mit drei Unbekannten vorgehen. Es sei

$$A = (a_{ij})_{i,j=1,2,3} \in M_3(K), \quad b = [b_1, b_2, b_3], \quad x = [x_1, x_2, x_3].$$

Um

$$Ax = b$$

zu lösen, wählen wir x_3 fest und betrachten die ersten beiden Gleichungen als Gleichungen in x_1 und x_2 :

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 &= b_1 - a_{13}x_3 \\ a_{21}x_1 + a_{22}x_2 &= b_2 - a_{23}x_3. \end{aligned}$$

Ist $\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} \neq 0$, so gibt es genau eine Lösung $[x_1, x_2]$, nämlich

$$x_1 = \frac{\begin{vmatrix} b_1 - a_{13}x_3 & a_{12} \\ b_2 - a_{23}x_3 & a_{22} \end{vmatrix}}{\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}} = \frac{\begin{vmatrix} b_1 & a_{12} \\ b_2 & a_{22} \end{vmatrix} - \begin{vmatrix} a_{13} & a_{12} \\ a_{23} & a_{22} \end{vmatrix} x_3}{\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}},$$

$$x_2 = \frac{\begin{vmatrix} a_{11} & b_1 \\ a_{21} & b_2 \end{vmatrix} - \begin{vmatrix} a_{11} & a_{13} \\ a_{21} & a_{23} \end{vmatrix} x_3}{\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}}.$$

Setzt man dies in die dritte Gleichung

$$a_{31}x_1 + a_{32}x_2 + a_{33}x_3 = b_3$$

ein und multipliziert mit $\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}$, so erhält man

$$\begin{aligned}
& a_{31} \begin{vmatrix} b_1 & a_{12} \\ b_2 & a_{22} \end{vmatrix} - a_{31} \begin{vmatrix} a_{13} & a_{12} \\ a_{23} & a_{22} \end{vmatrix} x_3 + a_{32} \begin{vmatrix} a_{11} & b_1 \\ a_{21} & b_2 \end{vmatrix} - a_{32} \begin{vmatrix} a_{11} & a_{13} \\ a_{21} & a_{23} \end{vmatrix} x_3 \\
& + a_{33} \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} x_3 = b_3 \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix},
\end{aligned}$$

also

$$\begin{aligned}
& (a_{31} \begin{vmatrix} a_{12} & a_{13} \\ a_{22} & a_{23} \end{vmatrix} - a_{32} \begin{vmatrix} a_{11} & a_{13} \\ a_{21} & a_{23} \end{vmatrix} + a_{33} \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}) x_3 \\
& = a_{31} \begin{vmatrix} a_{12} & b_1 \\ a_{22} & b_2 \end{vmatrix} - a_{32} \begin{vmatrix} a_{11} & b_1 \\ a_{21} & b_2 \end{vmatrix} + b_3 \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}.
\end{aligned}$$

Definiert man nun die Determinante von A als

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

so gilt also

$$x_3 = \frac{\begin{vmatrix} a_{11} & a_{12} & b_1 \\ a_{21} & a_{22} & b_2 \\ a_{31} & a_{32} & b_3 \end{vmatrix}}{\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}} = \frac{\det(a^1, a^2, b)}{\det(a^1, a^2, a^3)},$$

falls $\det A \neq 0$ ist. Bevor wir uns in Details verlieren, stoppen wir hier und führen auf induktivem Wege die Determinante einer $n \times n$ -Matrix ein.

Definition 15.2 Es sei R ein kommutativer Ring mit Eins und $n \in \mathbb{N}$, $n > 0$. $A = (a_{ij}) \in M_n(R)$ sei eine $n \times n$ -Matrix über R .

Ist $n = 1$, so sei $\det A := a_{11}$.

Ist $n > 1$, so sei

$$\begin{aligned}
\det A & := a_{11} \det A_{11} - a_{12} \det A_{12} + \dots \\
& \quad + (-1)^{k+1} a_{1k} \det A_{1k} + \dots \\
& \quad + (-1)^{n+1} a_{1n} \det A_{1n}.
\end{aligned}$$

Hier ist A_{1k} die Untermatrix von A , die durch Streichen der ersten Zeile und der k -ten Spalte entsteht.

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

Man sieht sofort

Satz 15.3 Ist A eine untere Dreiecksmatrix,

$$A = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ * & & \lambda_n \end{pmatrix}, \quad \text{so gilt: } \det A = \lambda_1 \lambda_2 \cdot \dots \cdot \lambda_n.$$

Beweis: Nach Definition ist

$$\det A = \lambda_1 \det A', \quad \text{wobei} \quad A' = \begin{pmatrix} \lambda_2 & & 0 \\ & \ddots & \\ * & & \lambda_n \end{pmatrix}.$$

Die Behauptung folgt also durch Induktion nach n . □

Wir fassen jetzt $A \mapsto \det A$ als Funktion der Spalten a^1, \dots, a^n von A auf.

Es gilt:

$$\begin{aligned} R^{n,n} &= R^n \times \dots \times R^n = \{(a^1, \dots, a^n) \mid a^i \in R^n\} \\ &\quad \text{und} \\ \det &: R^n \times \dots \times R^n \longrightarrow R \end{aligned}$$

ist also eine Funktion in n Veränderlichen; jede Veränderliche ist eine Spalte $a = [a_1, \dots, a_n] \in R^n$. In R^n kann man genauso rechnen wie in K^n . Allgemein haben wir für beliebiges $m \in \mathbb{N}$, $m > 0$, die folgende

Definition 15.4 Eine Funktion $\varphi : R^n \times \dots \times R^n \rightarrow R$ in m Veränderlichen $a^1, \dots, a^m \in R^n$ heißt

- a) **m -fach multilinear** \Leftrightarrow φ ist in jeder Variablen linear, d.h. für $i = 1, \dots, m$ gilt:

$$\begin{aligned} &\varphi(a^1, \dots, a^{i-1}, \alpha a + \beta b, a^{i+1}, \dots, a^m) \\ &= \alpha \varphi(a^1, \dots, a^{i-1}, a, a^{i+1}, \dots, a^m) \\ &\quad + \beta \varphi(a^1, \dots, a^{i-1}, b, a^{i+1}, \dots, a^m) \end{aligned}$$

für alle $a^1, \dots, a^{i-1}, a, b, a^{i+1}, \dots, a^m \in R^n$, $\alpha, \beta \in R$.

- b) **alternierend** \Leftrightarrow Für alle $a^1, \dots, a^m \in R^n$ gilt:

$$\varphi(a^1, \dots, a^m) = 0, \quad \text{falls } a^i = a^{i+1} \text{ für einen Index } i \text{ mit } 1 \leq i < m.$$

Eine m -fach multilineare, alternierende Funktion $\varphi : R^n \times \dots \times R^n \rightarrow R$ heißt kurz eine **alternierende m -Form** auf R^n .

Die fundamentalen Eigenschaften der Determinante beschreibt

Satz 15.5 $\det : R^n \times \dots \times R^n \rightarrow R$ ist eine alternierende n -Form auf R^n mit $\det(e^1, \dots, e^n) = 1$.

Ausführlich heißt das:

- (1) $A \mapsto \det A$ ist linear in jeder Spalte.
- (2) $\det A = 0$, falls zwei benachbarte Spalten gleich sind.
- (3) $\det E_n = 1$.

Beweis durch Induktion nach n :

Für $n = 1$ sind (1) und (3) erfüllt, weil $\det = id_R$; die Aussage (2) ist für $n = 1$ leer, also auch erfüllt.

Induktionsschluß: $n - 1 \rightarrow n, n \geq 2$:

Zu (1): Es seien $a^1, \dots, a^{i-1}, a^{i+1}, \dots, a^n \in R^n$ fest, $a^k = [a_{1k}, \dots, a_{nk}]$ ($1 \leq k \leq n, k \neq i$). Es sei $\varphi(x) := \det(a^1, \dots, a^{i-1}, x, a^{i+1}, \dots, a^n)$ für $x = [x_1, \dots, x_n] \in R^n$. Für $\tilde{x} = [x_2, \dots, x_n]$, $j = 1, \dots, n, j \neq i$ sei

$$\begin{aligned}\varphi_j(\tilde{x}) &:= \det(\tilde{a}^1, \dots, \tilde{a}^{j-1}, \tilde{a}^{j+1}, \dots, \tilde{a}^{i-1}, \tilde{x}, \tilde{a}^{i+1}, \dots, \tilde{a}^n) \\ &= \det A_{1j},\end{aligned}$$

wobei $A = (a^1, \dots, a^{i-1}, x, a^{i+1}, \dots, a^n)$ und $\tilde{a}^k = [a_{2k}, \dots, a_{nk}]$ ($1 \leq k \leq n, k \neq i$).

Nach Definition der Determinante gilt nun:

$$\begin{aligned}\varphi(x) &= a_{11}\varphi_1(\tilde{x}) - a_{12}\varphi_2(\tilde{x}) + \dots + (-1)^i a_{1,i-1}\varphi_{i-1}(\tilde{x}) \\ &\quad + (-1)^{i+1} x_1 \det A_{1i} + \dots + (-1)^{n+1} a_{1n}\varphi_n(\tilde{x}).\end{aligned}$$

Nach Induktionsvoraussetzung sind $\varphi_1, \dots, \varphi_{i-1}, \varphi_{i+1}, \dots, \varphi_n$ linear in \tilde{x} , und somit ist φ linear in x .

Zu (2): Mit denselben Notationen wie im Beweis zu (1) ergibt sich

$$\begin{aligned}\det(a^1, \dots, a^{i-1}, x, x, a^{i+2}, \dots, a^n) &= a_{11} \det(\tilde{a}^2, \dots, \tilde{a}^{i-1}, \tilde{x}, \tilde{x}, \dots, \tilde{a}^n) \\ &\quad + \dots + (-1)^{i+1} x_1 \det(\tilde{a}^1, \dots, \tilde{a}^{i-1}, \tilde{x}, \tilde{a}^{i+2}, \dots, \tilde{a}^n) \\ &\quad + \dots + (-1)^{i+2} x_1 \det(\tilde{a}^1, \dots, \tilde{a}^{i-1}, \tilde{x}, \tilde{a}^{i+2}, \dots, \tilde{a}^n) \\ &\quad + \dots + (-1)^{n+1} a_{1n} \det(\tilde{a}^1, \dots, \tilde{x}, \tilde{x}, \dots, \tilde{a}^{n-1}) = 0.\end{aligned}$$

Zu (3): Das ist klar. □

Wir leiten nun einige Folgerungen aus (1) und (2) ab. Man beachte, dass diese Folgerungen dann sinngemäß auch für jede alternierende m -Form auf R^n gelten.

Zunächst kann man (1) ausführlich schreiben als

$$\det(a^1, \dots, a^{i-1}, \lambda a^i, a^{i+1}, \dots, a^n) = \lambda \det(a^1, \dots, a^n); \quad (1')$$

$$\begin{aligned}\det(a^1, \dots, a^{i-1}, a + b, a^{i+1}, \dots, a^n) &= \det(a^1, \dots, a^{i-1}, a, a^{i+1}, \dots, a^n) \\ &\quad + \det(a^1, \dots, a^{i-1}, b, a^{i+1}, \dots, a^n)\end{aligned} \quad (1'')$$

für alle $a^1, \dots, a^n, a, b \in R^n, \lambda \in R$. Insbesondere gilt für alle $\lambda \in R, A \in M_n(R)$:

$$(4) \quad \det(\lambda A) = \lambda^n \det A.$$

Weiter gilt

Lemma 15.6 Es sei $A \in M_n(R)$. Dann gilt:

(5) $\det A = 0$, falls eine Spalte von A Null ist.

(6) $\det B = -\det A$, falls B aus A durch Vertauschen zweier Spalten entsteht.

(7) $\det A = 0$, falls zwei Spalten von A gleich sind.

Beweis zu (5): Es sei $a^i = 0$.

$$\varphi(x) = \det(a^1, \dots, a^{i-1}, x, a^{i+1}, \dots, a^n)$$

ist eine lineare Funktion $\varphi : R^n \rightarrow R$, also ist $\varphi(0) = 0$. Wegen $a^i = 0$ ist daher $\det A = \varphi(0) = 0$.

Zu (6):

a) Zwei benachbarte Spalten werden vertauscht:

$$B = (a^1, \dots, a^{i-1}, a^{i+1}, a^i, a^{i+2}, \dots, a^n).$$

Sei

$$\varphi(x, y) := \det(a^1, \dots, a^{i-1}, x, y, a^{i+2}, \dots, a^n).$$

Dann gilt:

$$\begin{aligned} 0 &\stackrel{(2)}{=} \varphi(x+y, x+y) \stackrel{(1)}{=} \varphi(x, x) + \varphi(x, y) + \varphi(y, x) + \varphi(y, y) \\ &\stackrel{(2)}{=} \varphi(x, y) + \varphi(y, x), \end{aligned}$$

also

$$\det B = \varphi(a^{i+1}, a^i) = -\varphi(a^i, a^{i+1}) = -\det A.$$

b) Es sei $k \geq 1$ und $1 \leq i \leq n - k$:

$$\varphi(x^1, \dots, x^{k+1}) = \det(a^1, \dots, a^{i-1}, x^1, \dots, x^{k+1}, a^{i+k+1}, \dots, a^n).$$

Wir müssen zeigen dass,

$$\varphi(x^1, \dots, x^{k+1}) = -\varphi(x^{k+1}, x^2, \dots, x^k, x^1)$$

gilt. Das geht durch Induktion nach k ($1 \leq k \leq n - i$). Für $k = 1$ liegt der Fall a) vor, den wir schon bewiesen haben.

Induktionsschluß: $k - 1 \rightarrow k$, ($2 \leq k \leq n - i$):

$$\begin{aligned} \varphi(x^1, \dots, x^{k+1}) &= -\varphi(x^2, x^1, \dots, x^{k+1}) = \varphi(x^2, x^{k+1}, x^3, \dots, x^k, x^1) \\ &= -\varphi(x^{k+1}, x^2, \dots, x^k, x^1). \end{aligned}$$

Zu (7): $\det A = \det(\dots, a, \dots, a, \dots) \stackrel{(6)}{=} -\det(\dots, a, \dots, a, \dots) = -\det A$
 $\Rightarrow 2 \det A = 0 \Rightarrow \det A = 0$, falls $2 \cdot 1_R$ ein Nichtnullteiler in R ist. Ist $2 \cdot 1_R$ Nullteiler in R , so kann man nicht so vorgehen.

Deshalb geht man besser folgendermaßen vor: Nach (2) gilt $\det A = 0$, falls zwei benachbarte Spalten gleich sind. Sind nun zwei nicht benachbarte Spalten gleich, so gilt:

$$\det A = \det(\dots, a, b, \dots, a, \dots) \stackrel{(6)}{=} -\det(\dots, a, a, \dots, b, \dots) \stackrel{(2)}{=} 0.$$

□

Lemma 15.7 Die Determinante von A bleibt unverändert, wenn man das λ -fache der j -ten Spalte zur i -ten Spalte addiert (wobei $\lambda \in R$ und $i \neq j$).

Beweis:

$$\begin{aligned} \det(a^1, \dots, a^i + \lambda a^j, \dots, a^j, \dots, a^n) &= \det(a^1, \dots, a^n) \\ &\quad + \lambda \det(a^1, \dots, a^j, \dots, a^j, \dots, a^n) \\ &\stackrel{(7)}{=} \det(a^1, \dots, a^n). \end{aligned}$$

□

Bemerkung 15.8 Man kann die Determinante $\det A$ einer $n \times n$ -Matrix über einem Körper K mit Hilfe des Gaußschen Eliminationsverfahrens berechnen: Nach (6) ändert die Determinante von A nach Multiplikation von P_{ij} von rechts das Vorzeichen, und nach 15.7 bleibt die Determinante bei Multiplikation mit $Q_{ij}(\lambda)$, $\lambda \in K$, unverändert.

Beispiel 15.9

a)

$$\begin{aligned} \det \begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 0 \\ 1 & 0 & 0 \end{pmatrix} &= -\det \begin{pmatrix} 1 & 1 & 1 \\ 0 & 2 & 1 \\ 0 & 0 & 1 \end{pmatrix} = -\det \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 1 \end{pmatrix} \\ &= -\det \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{pmatrix} = -2. \end{aligned}$$

b)

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

Allgemein kann man $A \in M_n(K)$ durch elementare Spaltenoperationen vom Typ P_{ij} , $Q_{ij}(\lambda)$ auf eine Spaltenstufenform bringen. Ist $\operatorname{rg} A < n$, so kommt eine Nullspalte vor, und somit ist dann nach Lemma 15.6 (5) $\det A = 0$. Ist aber $\operatorname{rg} A = n$, so hat die Spaltenstufenform die Gestalt einer unteren Dreiecksmatrix

$$\begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ * & \lambda_2 & & \vdots \\ \vdots & & \ddots & \vdots \\ * & \dots & & \lambda_n \end{pmatrix} \quad \text{mit } \lambda_i \in K, \lambda_i \neq 0.$$

Es gilt dann: $\det A = (-1)^k \lambda_1 \cdot \dots \cdot \lambda_n \neq 0$, wobei k die Anzahl der durchgeführten Spaltenvertauschungen ist. Insbesondere gilt

Satz 15.10 Ist K ein Körper und $A \in M_n(K)$, so sind folgende Aussagen äquivalent:

- a) $\det A \neq 0$.
 b) $\operatorname{rg} A = n$.
 c) A ist invertierbar.

Sei jetzt wieder R ein beliebiger kommutativer Ring mit Eins. Wir beweisen den fundamentalen

Satz 15.11 Es sei $D : R^n \times \dots \times R^n \rightarrow R$ eine alternierende n -Form auf R^n . Dann gilt:

$$D = D(e^1, \dots, e^n) \det,$$

d.h. $\forall a^1, \dots, a^n \in R^n$:

$$D(a^1, \dots, a^n) = D(e^1, \dots, e^n) \det(a^1, \dots, a^n).$$

Insbesondere ist \det die einzige alternierende n -Form auf R^n mit $\det(e^1, \dots, e^n) = 1$.

Beweis: Es sei

$$a^j = [a_{1j}, \dots, a_{nj}] = \sum_{i=1}^n a_{ij} e^i.$$

Dann gilt nach den Regeln (1), (2) und (7), die ja auch für D gelten:

$$\begin{aligned} D(a^1, \dots, a^n) &= \sum_{i_1=1}^n a_{i_1 1} D(e^{i_1}, a^2, \dots, a^n) = \dots \\ &= \sum_{i_1=1}^n \dots \sum_{i_n=1}^n a_{i_1 1} a_{i_2 2} \dots a_{i_n n} D(e^{i_1}, \dots, e^{i_n}) \\ &= \sum_{(i_1, \dots, i_n)} a_{i_1 1} \dots a_{i_n n} D(e^{i_1}, \dots, e^{i_n}), \end{aligned}$$

wobei über alle Permutationen (i_1, \dots, i_n) von $(1, \dots, n)$ summiert wird. (Eine **Permutation** von $(1, \dots, n)$ ist ein n -Tupel (i_1, \dots, i_n) von Zahlen $i_\nu \in \{1, \dots, n\}$ mit der Eigenschaft $i_\nu \neq i_\mu$, falls $\nu \neq \mu$.) Analog ergibt sich

$$\det(a^1, \dots, a^n) = \sum_{(i_1, \dots, i_n)} a_{i_1 1} \dots a_{i_n n} \det(e^{i_1}, \dots, e^{i_n}).$$

Wir müssen also nur noch zeigen, dass

$$D(e^{i_1}, \dots, e^{i_n}) = D(e^1, \dots, e^n) \det(e^{i_1}, \dots, e^{i_n}) \quad (*)$$

gilt. Das ist aber wegen der Eigenschaft (6), die sowohl für D als auch für \det gilt, einfach: Seien k Vertauschungen erforderlich, um (i_1, \dots, i_n) in die natürliche Reihenfolge $(1, \dots, n)$ zu bringen. Dann gilt nach (6):

$$D(e^{i_1}, \dots, e^{i_n}) = (-1)^k D(e^1, \dots, e^n)$$

und

$$\det(e^{i_1}, \dots, e^{i_n}) = (-1)^k \det(e^1, \dots, e^n) = (-1)^k.$$

Also gilt (*). □

Bemerkung 15.12 Ist K ein Körper, so ist

$$A_m(K^n) = \{\varphi \mid \varphi \text{ alternierende } m\text{-Form auf } K^n\}$$

ein K -Untervektorraum von Abb $(\underbrace{K^n \times \dots \times K^n}_{m\text{-mal}}, K)$.

Satz 15.11 besagt, dass $A_n(K^n)$ eindimensional ist. Jedes Element $D \in A_n(K^n)$, $D \neq 0$, (z.B. $D = \det$) ist daher eine Basis von $A_n(K^n)$:

$$A_n(K^n) = \{\lambda D \mid \lambda \in K\}.$$

Es sei R ein kommutativer Ring mit Eins. Es gilt der folgende Determinantenmultiplikationssatz.

Satz 15.13 Für $A, B \in M_n(R)$ gilt:

$$\det(AB) = \det A \cdot \det B.$$

Beweis: Es sei $A \in M_n(R)$ gegeben. $D(C) := \det(AC)$ ist alternierende n -Form in den Spalten von C , wie man leicht sieht, weil ja $x \mapsto Ax$ linear ist und $D(x^1, \dots, x^n) = \det(Ax^1, \dots, Ax^n)$. Nach Satz 15.11 gilt:

$$D = D(E) \det = (\det A) \det,$$

also

$$\det(AB) = D(B) = D(E) \det B = \det A \det B.$$

□

Definition 15.14 Es sei S ein beliebiger Ring mit Eins. $a \in S$ heißt **Einheit** in S (oder: **invertierbar** in S), wenn gilt: $\exists b \in S$ mit $ab = ba = 1_S$. b ist eindeutig bestimmt, denn:

$$ab' = b'a = 1_S \quad \Rightarrow \quad b' = 1_S b' = (ba)b' = b(ab') = b1_S = b.$$

b heißt das **Inverse** von a und wird mit a^{-1} bezeichnet.

Mit S^\times wird die Menge aller Einheiten in S bezeichnet. Sind $a, b \in S^\times$, so ist auch $ab \in S^\times$ mit $(ab)^{-1} = b^{-1}a^{-1}$.

Als Folgerung aus dem Determinantenmultiplikationssatz erhalten wir

Korollar 15.15 Ist $A \in M_n(R)$ invertierbar, so ist $\det A$ eine Einheit in R .

Beweis: $AB = E_n$ impliziert nach Satz 15.13 $\det A \det B = 1$. □

Die Umkehrung gilt auch, wird aber erst etwas später bewiesen.

Satz 15.16 Für alle $A \in M_n(R)$ gilt:

$$\det A = \det A^t.$$

Beweis: Es sei $A = (a_{ij})$. Für eine Permutation (i_1, \dots, i_n) von $(1, \dots, n)$ sei mit (j_1, \dots, j_n) die inverse Permutation bezeichnet. Dann gilt also $i_{j_\nu} = \nu$ für $\nu = 1, \dots, n$; und somit

$$a_{i_1 1} \cdot \dots \cdot a_{i_n n} = a_{1 j_1} \cdot \dots \cdot a_{n j_n}$$

und wegen $(e^{i_1}, \dots, e^{i_n})e^{j_\nu} = e^{i_{j_\nu}} = e^\nu$ auch

$$(e^{i_1}, \dots, e^{i_n})(e^{j_1}, \dots, e^{j_n}) = (e^1, \dots, e^n).$$

Nach Satz 15.13 gilt also:

$$\det(e^{i_1}, \dots, e^{i_n}) \det(e^{j_1}, \dots, e^{j_n}) = \det(e^1, \dots, e^n) = 1.$$

Da diese Determinanten nur die Werte ± 1 annehmen, folgt somit

$$\det(e^{i_1}, \dots, e^{i_n}) = \det(e^{j_1}, \dots, e^{j_n}).$$

Man erhält also

$$\begin{aligned} \det A &= \sum_{(i_1, \dots, i_n)} a_{i_1 1} \cdot \dots \cdot a_{i_n n} \det(e^{i_1}, \dots, e^{i_n}) \\ &= \sum_{(j_1, \dots, j_n)} a_{1 j_1} \cdot \dots \cdot a_{n j_n} \det(e^{j_1}, \dots, e^{j_n}) = \det A^t. \end{aligned}$$

□

Korollar 15.17 Die Determinantenfunktion $A \mapsto \det A$ ist auch multilinear und alternierend als Funktion der Zeilen von A .

Definition 15.18 Es sei $A = (a_{ij}) \in M_n(R)$. Mit A_{ij} wird die Matrix bezeichnet, die man durch Streichen der i -ten Zeile und der j -ten Spalte aus A erhält.

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

heißt das **algebraische Komplement** von a_{ij} in A (oder auch der **Kofaktor** von a_{ij} in A).

$$\tilde{A} = (\tilde{a}_{ji})_{\substack{j=1, \dots, n \\ i=1, \dots, n}} \in M_n(R)$$

heißt die zu A **komplementäre Matrix**.

Beispiel 15.19

$$\text{Für } A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \text{ ist } \tilde{A} = \begin{pmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{pmatrix}.$$

Es gilt nun der wichtige

Satz 15.20 Sei $A = (a_{ij}) \in M_n(R)$. Dann gilt:

$$\text{a) } \det A \cdot \delta_{ki} = \sum_{j=1}^n a_{kj} \tilde{a}_{ji} \text{ für alle } k, i.$$

$$\text{b) } \det A \cdot \delta_{jk} = \sum_{i=1}^n \tilde{a}_{ji} a_{ik} \text{ f\"ur alle } j, k.$$

In Matrixschreibweise:

$$(\det A)E_n = A\tilde{A} = \tilde{A}A.$$

Beweis: zu b):

$$\begin{aligned} \tilde{a}_{ji} &= (-1)^{i+j} \det \begin{pmatrix} 1 & 0 \\ 0 & A_{ij} \end{pmatrix} \\ &= \det \begin{pmatrix} a_{11} & \dots & 0 & \dots & a_{1n} \\ \vdots & & \vdots & & \vdots \\ 0 & \dots & 1 & \dots & 0 \\ \vdots & & \vdots & & \vdots \\ a_{n1} & \dots & 0 & \dots & a_{nn} \end{pmatrix} = \det(a^1, \dots, a^{j-1}, e^i, a^{j+1}, \dots, a^n), \end{aligned}$$

wobei die Eins in der i -ten Zeile und j -ten Spalte steht.

Es folgt

$$\sum_{i=1}^n \tilde{a}_{ji} a_{ik} = \det(a^1, \dots, a^{j-1}, a^k, a^{j+1}, \dots, a^n) = \begin{cases} \det A, & k = j \\ 0, & k \neq j \end{cases}.$$

Es gilt also: $\tilde{A}A = (\det A)E_n$.

Zu a): Das zeigt man genauso oder man geht so vor: Nach Satz 15.16 gilt $\tilde{A}^t = (\tilde{A})^t$, und somit gilt:

$$(A\tilde{A})^t = (\tilde{A})^t A^t = \tilde{A}^t A^t = (\det A^t)E_n = (\det A)E_n.$$

Transponiert man diese Gleichung, so ergibt sich $A\tilde{A} = (\det A)E_n$. □

Korollar 15.21 Ist $A \in M_n(R)$ und $\det A$ eine Einheit in R , so ist A invertierbar, und es gilt:

$$A^{-1} = \frac{1}{\det A} \tilde{A}.$$

Diese Formel ist mehr von theoretischem Interesse und nicht so sehr zur Berechnung von A^{-1} geeignet. Dasselbe gilt f\"ur das folgende

Korollar 15.22 (Cramersche Regel) Ist $A \in M_n(R)$ invertierbar, $b \in R^n$, so besitzt das lineare Gleichungssystem

$$Ax = b$$

die eindeutig bestimmte L\"osung

$$x = A^{-1}b = \frac{1}{\det A} \tilde{A}b.$$

Für die j -te Komponente x_j von $x \in R^n$ gilt also:

$$\begin{aligned} x_j &= \frac{1}{\det A} \sum_{i=1}^n \tilde{a}_{ji} b_i = \frac{1}{\det A} \sum_{i=1}^n \det(a^1, \dots, a^{j-1}, e^i, a^{j+1}, \dots, a^n) b_i \\ &= \frac{\det(a^1, \dots, a^{j-1}, b, a^{j+1}, \dots, a^n)}{\det(a^1, \dots, a^n)}. \end{aligned}$$

Übungen

1. Sei $A = (a_{ij}) \in M_n(R)$. Dann gilt:

$$\det A = \sum_{i=1}^n a_{ik} (-1)^{i+k} \det A_{ik}$$

für $k = 1, \dots, n$ (**Entwicklung nach der k -ten Spalte**),

$$\det A = \sum_{i=1}^n a_{ki} (-1)^{i+k} \det A_{ki}$$

für $k = 1, \dots, n$ (**Entwicklung nach der k -ten Zeile**).

2. Beweisen Sie:

(a)

$$\Delta(x_1, \dots, x_n) := \det \begin{pmatrix} 1 & x_1 & x_1^2 & \dots & x_1^{n-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_n & x_n^2 & \dots & x_n^{n-1} \end{pmatrix} = \prod_{1 \leq i < j \leq n} (x_j - x_i).$$

(b) Stellen Sie $\Delta(x_1, x_2, x_3)^2$ als Funktion von a_1, a_2, a_3 dar, wobei

$$(X - x_1)(X - x_2)(X - x_3) = X^3 - a_1 X^2 + a_2 X - a_3.$$

3. Für $A \in M_n(R)$, $B \in R^{n,m}$, $C \in M_m(R)$ ist

$$\det \begin{pmatrix} A & B \\ 0 & C \end{pmatrix} = \det A \det C$$

und

$$\det \begin{pmatrix} B & A \\ C & 0 \end{pmatrix} = (-1)^{nm} \det A \det C.$$

4. Es gilt $1798 = 31 \cdot 58$, $2139 = 31 \cdot 69$, $3255 = 31 \cdot 105$ und $4867 = 31 \cdot 157$. Beweisen Sie ohne irgendwelche Berechnungen, dass 31 ein Teiler der Determinante

$$\det \begin{pmatrix} 1 & 7 & 9 & 8 \\ 2 & 1 & 3 & 9 \\ 3 & 2 & 5 & 5 \\ 4 & 8 & 6 & 7 \end{pmatrix}$$

ist.

5. Es sei R ein kommutativer Ring mit Eins. Beweisen Sie:

- (a) $\widetilde{AB} = \widetilde{B} \widetilde{A}$ für $A, B \in M_n(R)$.
- (b) $\widetilde{(aA)} = a^{n-1} \widetilde{A}$ für $a \in R, A \in M_n(R)$.
- (c) $\widetilde{\widetilde{A}} = (\det A)^{n-2} A$ für $A \in M_n(R)$.
- (d) $\det \widetilde{A} = (\det A)^{n-1}$ für $A \in M_n(R)$.

6. Für $a^1, \dots, a^{n-1} \in R^n$ sei $a^1 \times \dots \times a^{n-1} \in R^n$ das n -Tupel $[d_1, \dots, d_n]$ mit

$$d_i = (-1)^{i+1} \det A_i,$$

wobei $A := (a^1, \dots, a^{n-1}) \in R^{n, n-1}$ und A_i aus A durch Streichen der i -ten Zeile entsteht.

(a) Beweisen Sie: Für $b = [b_1, \dots, b_n] \in R^n$ gilt:

$$\det(b, a^1, \dots, a^{n-1}) = (a^1 \times \dots \times a^{n-1})^t b.$$

(b) Es sei K ein Körper und $A \in K^{n-1, n}$ eine Matrix vom Rang $n-1$. a^1, \dots, a^{n-1} seien die Spalten von A^t . Zeigen Sie: Das lineare Gleichungssystem

$$Ax = 0$$

hat die Lösungsmenge

$$K(a^1 \times \dots \times a^{n-1}).$$

7. Es sei K ein Körper und $A \in K^{r, n}$ eine $r \times n$ -Matrix vom Rang $r \leq n$. Es sei $B \in K^{n-r, n}$, so dass $C = \begin{pmatrix} A \\ B \end{pmatrix} \in M_n(K)$ invertierbar ist. Zeigen Sie: Die letzten $n-r$ Spalten von \widetilde{C} bilden eine Basis von $\{x \in K^n \mid Ax = 0\}$.

8. Es sei $A \in M_n(K)$. Für $x^1, \dots, x^n \in K^n$ sei

$$\varphi(x^1, \dots, x^n) := \sum_{j=1}^n \det(x^1, \dots, x^{j-1}, Ax^j, x^{j+1}, \dots, x^n).$$

- (a) Zeigen Sie: $\varphi \in A_n(K^n)$.
- (b) Berechnen Sie $\lambda \in K$ mit $\varphi = \lambda \det$. λ heißt die **Spur** von A .
- (c) Ist $K = \mathbb{R}$, so ist $f(t) := \det(E + At)$ ein reelles Polynom in t . Berechnen Sie die Ableitung $f'(0)$.

9. Es sei V ein n -dimensionaler K -Vektorraum. Alternierende m -Formen auf V sind wie in Definition 15.4 erklärt.

(a) Beweisen Sie:

$$A_n(V) = \{\varphi \mid \varphi \text{ alternierende } n\text{-Form auf } V\}$$

ist ein eindimensionaler K -Vektorraum.

(b) $F \in \text{End}(V)$ induziert eine lineare Abbildung

$$F^* : A_n(V) \longrightarrow A_n(V)$$

mit

$$(F^*(\varphi))(v_1, \dots, v_n) := \varphi(F(v_1), \dots, F(v_n))$$

für alle $v_1, \dots, v_n \in V$. Welche Bedeutung hat der Koeffizient $\lambda \in K$ mit $F^* = \lambda \text{id}_{A_n(V)}$?

(c) Sei $F \in \text{End}(V)$ und $\hat{F} : A_n(V) \rightarrow A_n(V)$ die Abbildung mit

$$(\hat{F}(\varphi))(v_1, \dots, v_n) := \sum_{i=1}^n \varphi(v_1, \dots, v_{i-1}, F(v_i), v_{i+1}, \dots, v_n).$$

Zeigen Sie: \hat{F} ist linear.

Der Koeffizient $\lambda \in K$ mit $\hat{F} = \lambda \text{id}$ heißt die **Spur** von F . Vergleichen Sie dies mit Aufgabe 8.

16 Eigenwerte, Eigenvektoren, charakteristisches Polynom

Es sei K ein Körper und V ein n -dimensionaler K -Vektorraum.

Wir wollen nun die Endomorphismen von V , d.h. die linearen Abbildungen

$$F : V \rightarrow V$$

von V in sich genauer untersuchen.

Ist $\mathcal{B} = (v_1, \dots, v_n)$ eine Basis von V , so gibt es zu $F \in \text{End}(V)$ genau eine Matrix $A = (a_{ij}) \in M_n(K)$, so dass

$$F(v_j) = \sum_{i=1}^n a_{ij} v_i$$

für $j = 1, \dots, n$ gilt.

A heißt die Matrix von F bezüglich der Basis \mathcal{B} .

$$\text{End}(V) \longrightarrow M_n(K), \quad F \longmapsto A$$

ist ein K -Algebra-Isomorphismus.

Ist $\mathcal{B}' = (w_1, \dots, w_n)$ eine andere Basis von V und gilt $v = \sum_{i=1}^n x_i v_i = \sum_{j=1}^n x'_j w_j$, so ist $x = [x_1, \dots, x_n]$ der Koordinatenvektor von v bzgl. \mathcal{B} und $x' = [x'_1, \dots, x'_n]$ der Koordinatenvektor von v bzgl. \mathcal{B}' . Ist nun $S = (s_{ij})$ die Basiswechselmatrix mit $v_i = \sum_{j=1}^n s_{ji} w_j$, so ist

$$x' = Sx,$$

und aus $y = Ax$, $y' = Sy$ folgt:

$$y' = SAx = (SAS^{-1})x'.$$

$B = SAS^{-1} \in M_n(K)$ ist also die Matrix von F bzgl. der Basis \mathcal{B}' .

Das Ziel ist es nun, für einen Endomorphismus

$$F : V \longrightarrow V$$

eine Basis $\mathcal{B} = (v_1, \dots, v_n)$ von V zu finden, so dass die Matrix A von F bzgl. \mathcal{B} eine besonders einfache Gestalt hat. \mathcal{B} wird dabei von dem gegebenen Endomorphismus abhängen.

Man kann dieses Problem auch so beschreiben: Zu gegebener Matrix $A \in M_n(K)$ finde man eine invertierbare Matrix $S \in M_n(K)$, so dass die Matrix SAS^{-1} eine besonders einfache Gestalt hat.

Wir beginnen unsere Untersuchung mit folgender

Definition 16.1 Es sei $F \in \text{End}(V)$. Ein Untervektorraum V_0 von V heißt **invariant** bezüglich des Endomorphismus F (kurz: *F-invariant*), wenn

$$F(V_0) \subset V_0$$

gilt.

Beispiel 16.2

a) Natürlich sind 0 und V trivialerweise F -invariant. Auch alle Untervektorräume von $\ker F$ sind invariant.

b) Wir betrachten

$$A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} : \mathbb{R}^2 \longrightarrow \mathbb{R}^2.$$

A besitzt außer 0 und \mathbb{R}^2 keine invarianten Unterräume. Ist nämlich $0 \neq V_0 \subset \mathbb{R}^2$ ein invarianter Unterraum, so gibt es ein $\begin{pmatrix} x \\ y \end{pmatrix} \in V_0 \setminus \{0\}$. Es gilt: $A \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} y \\ -x \end{pmatrix} \in V_0$. Da

$$\det \begin{pmatrix} x & y \\ y & -x \end{pmatrix} = -(x^2 + y^2) < 0$$

(weil $\begin{pmatrix} x \\ y \end{pmatrix} \neq 0$ ist), sind $\begin{pmatrix} x \\ y \end{pmatrix}$ und $A \begin{pmatrix} x \\ y \end{pmatrix}$ linear unabhängig, also $\dim V_0 = 2$, d.h. $V_0 = \mathbb{R}^2$.

c) Wir betrachten

$$A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} : \mathbb{C}^2 \longrightarrow \mathbb{C}^2.$$

Dann gilt:

$$A \begin{pmatrix} 1 \\ i \end{pmatrix} = \begin{pmatrix} i \\ -1 \end{pmatrix} = i \begin{pmatrix} 1 \\ i \end{pmatrix},$$

und somit ist $V_0 = \mathbb{C} \begin{pmatrix} 1 \\ i \end{pmatrix} \subset \mathbb{C}^2$ ein eindimensionaler invarianter Unterraum bzgl. A .

Auch $V_1 = \mathbb{C} \begin{pmatrix} 1 \\ -i \end{pmatrix}$ ist invariant, weil

$$A \begin{pmatrix} 1 \\ -i \end{pmatrix} = \begin{pmatrix} -i \\ -1 \end{pmatrix} = -i \begin{pmatrix} 1 \\ -i \end{pmatrix}$$

gilt.

Da $\begin{pmatrix} 1 \\ i \end{pmatrix}$ und $\begin{pmatrix} 1 \\ -i \end{pmatrix}$ \mathbb{C} -linear unabhängig sind, gilt:

$$\mathbb{C}^2 = V_0 \oplus V_1,$$

und A hat bzgl. der Basis $\begin{pmatrix} 1 \\ i \end{pmatrix}, \begin{pmatrix} 1 \\ -i \end{pmatrix}$ von \mathbb{C}^2 die Matrix

$$B = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix},$$

denn

$$A \begin{pmatrix} 1 \\ i \end{pmatrix} = i \begin{pmatrix} 1 \\ i \end{pmatrix} + 0 \begin{pmatrix} 1 \\ -i \end{pmatrix}$$

und

$$A \begin{pmatrix} 1 \\ -i \end{pmatrix} = 0 \begin{pmatrix} 1 \\ i \end{pmatrix} + (-i) \begin{pmatrix} 1 \\ -i \end{pmatrix}.$$

d)

$$A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} : \mathbb{C}^2 \longrightarrow \mathbb{C}^2$$

besitzt außer 0 und \mathbb{C}^2 noch genau einen eindimensionalen invarianten Unterraum, nämlich

$$V_0 = \ker A = \left\{ \begin{pmatrix} z \\ w \end{pmatrix} \in \mathbb{C}^2 \mid w = 0 \right\}.$$

Eine direkte Summenzerlegung von \mathbb{C}^2 in invariante Unterräume wie in Beispiel c) ist hier nicht möglich.

An folgendem einfachen Lemma sieht man, dass die F -invarianten Unterräume von V eine besondere Rolle für unser Problem, F durch eine besonders einfache Matrix zu beschreiben, spielen werden.

Lemma 16.3 Es sei $F \in \text{End}(V)$, und $V_0 \subset V$ sei ein k -dimensionaler F -invarianter Unterraum von V , $0 < k < n$. Es sei (v_1, \dots, v_k) eine Basis von V_0 , und v_{k+1}, \dots, v_n seien Vektoren in V , so dass $\mathcal{B} = (v_1, \dots, v_n)$ eine Basis von V ist. Dann hat die Matrix $A \in M_n(K)$ von F bzgl. \mathcal{B} die folgende Blockgestalt:

$$A = \begin{pmatrix} A' & B \\ 0 & A'' \end{pmatrix}$$

mit $A' \in M_k(K)$, $A'' \in M_{n-k}(K)$, $B \in K^{k, n-k}$.

Weiter gilt: Gibt es einen zu V_0 komplementären F -invarianten Unterraum V_1 , d.h. gilt $V = V_0 \oplus V_1$, so kann man (v_{k+1}, \dots, v_n) als Basis von V_1 wählen, und die Matrix A hat dann die Blockdiagonalgestalt

$$A = \begin{pmatrix} A' & 0 \\ 0 & A'' \end{pmatrix}.$$

Beweis:

a) Sei $A = (a_{ij})$. Dann ist

$$F(v_j) = \sum_{i=1}^n a_{ij} v_i \quad \text{für } j = 1, \dots, n.$$

Da $V_0 = \langle v_1, \dots, v_k \rangle$ invariant ist, gilt also:

$$F(v_j) \in \langle v_1, \dots, v_k \rangle \quad \text{für } j = 1, \dots, k;$$

also muss $a_{ij} = 0$ gelten, wenn $i > k$ und $j \leq k$ ist.

b) Ist auch $V_1 = \langle v_{k+1}, \dots, v_n \rangle$ invariant, so gilt außerdem: $a_{ij} = 0$, falls $j > k$ und $i \leq k$. \square

Die einfachste Klasse von Endomorphismen sind die diagonalisierbaren Endomorphismen.

Definition und Lemma 16.4 Es sei $F \in \text{End}(V)$. F heißt **diagonalisierbar** (über K), wenn eine der folgenden zueinander äquivalenten Bedingungen erfüllt ist:

- (1) V besitzt eine direkte Summenzerlegung in eindimensionale F -invariante Unterräume V_1, \dots, V_n :

$$V = V_1 \oplus \dots \oplus V_n, \quad \dim V_i = 1, \quad F(V_i) \subset V_i \quad \text{für } i = 1, \dots, n.$$

- (2) Es gibt eine Basis $\mathcal{B} = (v_1, \dots, v_n)$ von V , so dass die Matrix A von F bzgl. \mathcal{B} Diagonalgestalt besitzt:

$$A = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}, \quad F(v_i) = \lambda_i v_i \quad \text{für } i = 1, \dots, n.$$

Beweis: (1) \Rightarrow (2): Sei $v_i \in V_i \setminus \{0\}$. Dann ist $\mathcal{B} = (v_1, \dots, v_n)$ eine Basis von V und wegen $F(V_i) \subset V_i$, $V_i = K v_i$ folgt $F(v_i) = \lambda_i v_i$ mit $\lambda_i \in K$.

$$A = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$$

ist dann die Matrix von F bzgl. \mathcal{B} .

(2) \Rightarrow (1): Ist $\mathcal{B} = (v_1, \dots, v_n)$ Basis von V , so gilt $V = V_1 \oplus \dots \oplus V_n$, wobei $V_i := K v_i$. Es gilt $\dim V_i = 1$, und wegen $F(v_i) = \lambda_i v_i$ gilt $\forall a \in K$:

$$F(av_i) = aF(v_i) = a\lambda_i v_i \in V_i.$$

Also ist V_i ein F -invarianter Unterraum. □

Natürlich nennen wir eine Matrix $A \in M_n(K)$ **diagonalisierbar über K** , wenn der Endomorphismus

$$A : K^n \longrightarrow K^n, \quad x \longmapsto Ax$$

diagonalisierbar ist, und das bedeutet:

$\exists S \in M_n(K)$ invertierbar, so dass SAS^{-1} Diagonalmatrix ist.

Beispiel 16.5

$$A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \in \text{End}(\mathbb{R}^2)$$

ist nicht diagonalisierbar. Fasst man aber A als Endomorphismus von \mathbb{C}^2 auf, so ist A diagonalisierbar: Mit

$$S = \frac{1}{1+i} \begin{pmatrix} 1 & -i \\ 1 & i \end{pmatrix}$$

gilt:

$$SAS^{-1} = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}.$$

Die Matrix S hat die besondere Eigenschaft $\det S = 1$ und $S^{-1} = \overline{S}^t = \frac{i+1}{2} \begin{pmatrix} 1 & 1 \\ i & -i \end{pmatrix}$.

Der Endomorphismus

$$A : \mathbb{C}^2 \longrightarrow \mathbb{C}^2$$

hat bzgl. der Basis $\left(\begin{pmatrix} 1 \\ i \end{pmatrix}, \begin{pmatrix} 1 \\ -i \end{pmatrix}\right)$ von \mathbb{C}^2 die Matrix $\begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}$. A ist über \mathbb{C} , aber nicht über \mathbb{R} diagonalisierbar.

Um zu verstehen, wann ein Endomorphismus $F \in \text{End}(V)$ diagonalisierbar ist, müssen wir die eindimensionalen F -invarianten Untervektorräume von V untersuchen. Dazu benutzen wir

Definition 16.6 Es sei $F \in \text{End}(V)$.

- Ein Vektor $v \in V$ heißt **Eigenvektor** von F , wenn $v \neq 0$ ist und wenn der von v aufgespannte eindimensionale Untervektorraum Kv von V F -invariant ist, d.h. wenn es ein $\lambda \in K$ gibt, so dass $F(v) = \lambda v$ gilt.
- $\lambda \in K$ heißt **Eigenwert** von F , wenn es einen Vektor $v \in V$, $v \neq 0$ mit $F(v) = \lambda v$ gibt. v heißt dann Eigenvektor zum Eigenwert λ .
- Ist $\lambda \in K$ ein Eigenwert von F , so heißt $E(\lambda) = \ker(F - \lambda \text{id}_V)$ der **Eigenraum** von F zum Eigenwert λ . $E(\lambda)$ ist offensichtlich F -invarianter Untervektorraum von V mit $\dim E(\lambda) \geq 1$, und die Einschränkung von F auf $E(\lambda)$ ist das λ -fache der identischen Abbildung auf $E(\lambda)$.

$$\begin{array}{ccc} V & \xrightarrow{F} & V \\ \cup & & \cup \\ E(\lambda) & \xrightarrow{\lambda \text{id}} & E(\lambda) \end{array}$$

ist kommutativ.

Die Aussage 16.4 (2) können wir nun auch so formulieren:

F ist genau dann diagonalisierbar, wenn V eine Basis aus Eigenvektoren von F besitzt.

Satz 16.7 Es sei $F \in \text{End}(V)$. $\lambda_1, \dots, \lambda_m \in K$ seien paarweise verschiedene Eigenwerte von F , und $E_i = \ker(F - \lambda_i \text{id}_V)$, $i = 1, \dots, m$, seien die zugehörigen Eigenräume. Dann gilt:

- Ist v_i Eigenvektor von F zum Eigenwert λ_i , $i = 1, \dots, m$, so sind v_1, \dots, v_m linear unabhängig in V .
- Die Summe der Eigenräume E_1, \dots, E_m ist direkt, d.h.

$$\bigoplus_{i=1}^m E_i \subset V.$$

Insbesondere ist

$$\sum_{i=1}^m \dim E_i \leq n.$$

Beweis: zu a): Induktion nach m . Für $m = 1$ ist nichts zu beweisen. Induktionsschluß $m - 1 \rightarrow m$ ($m \geq 2$): Es seien $a_1, \dots, a_m \in K$ mit

$$a_1 v_1 + \dots + a_m v_m = 0. \quad (*)$$

Dann ist

$$\begin{aligned} \lambda_m a_1 v_1 + \dots + \lambda_m a_m v_m &= 0 = F(0) = F(a_1 v_1 + \dots + a_m v_m) \\ &= a_1 F(v_1) + \dots + a_m F(v_m) = a_1 \lambda_1 v_1 + \dots + a_m \lambda_m v_m. \end{aligned}$$

Also folgt

$$(\lambda_m - \lambda_1) a_1 v_1 + \dots + (\lambda_m - \lambda_{m-1}) a_{m-1} v_{m-1} = 0.$$

Nach Induktionsvoraussetzung folgt hieraus:

$$(\lambda_m - \lambda_i) a_i = 0 \quad \text{für } i = 1, \dots, m-1;$$

und somit $a_i = 0$ für $i = 1, \dots, m-1$. Aus (*) folgt dann $a_m v_m = 0$, also auch $a_m = 0$.

b): Es genügt zu zeigen, dass

$$\left(\sum_{i=1}^{m-1} E_i \right) \cap E_m = 0.$$

Es sei $v \in \left(\sum_{i=1}^{m-1} E_i \right) \cap E_m$, also $F(v) = \lambda_m v$ und

$$v = v_1 + \dots + v_{m-1} \quad \text{mit } F(v_i) = \lambda_i v_i.$$

Es folgt

$$F(v) = \sum_{i=1}^{m-1} F(v_i) = \sum_{i=1}^{m-1} \lambda_i v_i$$

und somit

$$\sum_{i=1}^{m-1} (\lambda_m - \lambda_i) v_i = 0.$$

Da $\lambda_m - \lambda_i \neq 0$ für $i = 1, \dots, m-1$, muss nach a) notwendig $v_1 = \dots = v_{m-1} = 0$, also auch $v = 0$ gelten. \square

Von fundamentaler Bedeutung ist nun die folgende

Definition 16.8

a) Es sei $A \in M_n(K)$, und t sei eine Unbestimmte. Dann ist

$$tE_n - A \in M_n(K[t])$$

eine Matrix mit Koeffizienten in dem Polynomring $K[t]$, und somit ist

$$P_A := \det(tE_n - A) \in K[t]$$

ein Polynom in der Unbestimmten t und Koeffizienten in K . P_A heißt das **charakteristische Polynom** von A .

- b) Es sei $F \in \text{End}(V)$ und $A \in M_n(K)$ die Matrix von F bzgl. einer Basis \mathcal{B} von V . Dann heißt $P_F := P_A$ das **charakteristische Polynom** von F .

Es gilt zunächst

Lemma 16.9

- a) Ist $A \in M_n(K)$, $B = SAS^{-1}$, wobei $S \in M_n(K)$ invertierbar ist, so gilt:

$$P_A = P_B.$$

- b) Das charakteristische Polynom P_F eines Endomorphismus $F \in \text{End}(V)$ ist unabhängig von der Wahl der Basis von V .

Beweis: zu a): Es gilt

$$tE_n - B = S(tE_n - A)S^{-1}.$$

Nach dem Determinantenmultiplikationssatz folgt hieraus:

$$P_B = (\det S)P_A(\det S)^{-1} = P_A.$$

- b) folgt unmittelbar aus a). □

Satz 16.10 Es sei $F \in \text{End}(V)$. Dann gilt:

- a) Das charakteristische Polynom $P_F \in K[t]$ hat die Gestalt

$$P_F = t^n - (\text{Spur } F)t^{n-1} + \dots + (-1)^n \det F.$$

Ist $A = (a_{ij}) \in M_n(K)$ die Matrix von F bzgl. einer Basis \mathcal{B} von V , so ist

$$\text{Spur } F := \text{Spur } A := \sum_{i=1}^n a_{ii}$$

und $\det F := \det A$. $\text{Spur } F$ heißt die **Spur** von F und $\det F$ heißt die **Determinante** von F .

- b) $\lambda \in K$ ist genau dann ein Eigenwert von F , wenn λ eine Nullstelle von P_F ist.

Beweis: zu a):

$$P_A = \det(tE_n - A) = (t - a_{11}) \cdot \dots \cdot (t - a_{nn}) + Q,$$

wobei $Q \in K[t]$ ein Polynom vom Grad $\leq n - 2$ ist. Also ist

$$P_A = t^n - (a_{11} + \dots + a_{nn})t^{n-1} + \text{Terme vom Grad } \leq n - 2.$$

Der konstante Term von P_A ist

$$P_A(0) = \det(-A) = (-1)^n \det A.$$

Zu b): Sei $\lambda \in K$. Dann gilt: $P_F(\lambda) = 0 \Leftrightarrow \det(\lambda E_n - A) = 0 \Leftrightarrow \text{rg}(\lambda E_n - A) < n \Leftrightarrow \lambda \text{id}_V - F$ ist nicht injektiv $\Leftrightarrow \exists v \in V, v \neq 0$, so dass $F(v) = \lambda v \Leftrightarrow \lambda$ ist Eigenwert von F . □

Bemerkung 16.11 Die Determinante und die Spur des Endomorphismus F sind **Invarianten** von F . Matrizen theoretisch drückt sich das so aus: Es gilt

$$\det A = \det B \quad \text{und} \quad \text{Spur} A = \text{Spur} B,$$

falls $B = SAS^{-1}$, S invertierbar. Ebenso sind natürlich auch die anderen Koeffizienten von P_F Invarianten von F .

Beispiel 16.12

a) Sei

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

und $K = \mathbb{Q}$. Dann ist

$$\begin{aligned} P_A &= \det \begin{pmatrix} t & 1 & -1 \\ 3 & t+2 & -3 \\ 2 & 2 & t-3 \end{pmatrix} = t \begin{vmatrix} t+2 & -3 \\ 2 & t-3 \end{vmatrix} - \begin{vmatrix} 3 & -3 \\ 2 & t-3 \end{vmatrix} - \begin{vmatrix} 3 & t+2 \\ 2 & 2 \end{vmatrix} \\ &= t(t+2)(t-3) + 6t - 3t + 9 - 6 - 6 + 2t + 4 \\ &= t^3 - t^2 - t + 1 = (t-1)^2(t+1). \end{aligned}$$

A besitzt die Eigenwerte 1 und -1 mit den Eigenräumen $E(1)$ und $E(-1)$.

Da $\text{rg}(A - E) = 1$ und $\text{rg}(A + E) = 2$ ist, folgt

$$\dim E(1) = 2, \quad \dim E(-1) = 1.$$

Nach Satz 16.7 muss also $\mathbb{Q}^3 = E(1) \oplus E(-1)$ gelten.

\mathbb{Q}^3 besitzt somit eine Basis aus Eigenvektoren von A . Also ist A diagonalisierbar über \mathbb{Q} .

b) Sei A dieselbe Matrix wie in a). Jetzt sei aber $K = \mathbb{F}_2$. Dann gilt (reduzieren modulo 2)

$$A \bmod 2 = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 0 & 0 & 1 \end{pmatrix},$$

und somit ist

$$P_{A \bmod 2} = \det \begin{pmatrix} t & 1 & 1 \\ 1 & t & 1 \\ 0 & 0 & t+1 \end{pmatrix} = (t+1) \det \begin{pmatrix} t & 1 \\ 1 & t \end{pmatrix} = (t+1)^3 \in \mathbb{F}_2[t].$$

(Natürlich ist $P_{A \bmod 2} = (P_A) \bmod 2$.)

Für den Eigenraum zum Eigenwert 1 gilt:

$$E(1) = \ker(A - E) = \ker \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 0 & 0 & 0 \end{pmatrix} \neq \mathbb{F}_2^3.$$

A ist somit über dem Körper \mathbb{F}_2 nicht diagonalisierbar.

- c) In den Beispielen a) und b) zerfiel das charakteristische Polynom P_A über dem Körper K in Linearfaktoren. Das ist aber im allgemeinen nicht der Fall: Sei $K = \mathbb{R}$ und

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in M_2(\mathbb{R}).$$

Dann ist

$$P_A = t^2 - (\text{Spur } A)t + \det A = t^2 - (a + d)t + ad - bc,$$

und A ist sicher nicht über \mathbb{R} diagonalisierbar, wenn das Polynom P_A keine reellen Nullstellen besitzt, d.h. wenn seine Diskriminante

$$\Delta = (\text{Spur } A)^2 - 4 \det A = (a - d)^2 + 4bc$$

negativ ist. A besitzt dann keine reellen Eigenwerte.

Ein Beispiel sind die von $\pm \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ verschiedenen Drehmatrizen

$$A = \begin{pmatrix} a & b \\ -b & a \end{pmatrix} \quad \text{mit } a, b \in \mathbb{R}, a^2 + b^2 = 1, b \neq 0.$$

(Hier ist $\Delta = -4b^2 < 0$.)

Für den Körper \mathbb{C} der komplexen Zahlen kann diese Komplikation nicht auftreten. Allgemein vereinbaren wir

Definition 16.13 Es sei K ein Körper. K heißt **algebraisch abgeschlossen**, wenn eine der folgenden äquivalenten Bedingungen erfüllt ist:

- Jedes nicht-konstante Polynom $P \in K[X]$ besitzt eine Nullstelle $\lambda \in K$.
- Jedes Polynom $P \in K[X]$ vom Grad $n \geq 1$ kann man in Linearfaktoren zerlegen, d.h. $\exists \alpha_1, \dots, \alpha_n \in K$ und $a \in K$, so dass

$$P = a(X - \alpha_1) \dots (X - \alpha_n).$$

Ohne Beweis notieren wir

Satz 16.14 Jeder Körper K besitzt einen Erweiterungskörper $\overline{K} \supset K$, der algebraisch abgeschlossen ist und dessen Elemente α algebraisch über K sind.

\overline{K} heißt ein **algebraischer Abschluss** von K .

Einen Beweis findet man z.B. bei Lang [18].

\mathbb{C} ist ein algebraischer Abschluss von \mathbb{R} .

Der **Fundamentalsatz der Algebra** lautet nämlich

Satz 16.15 \mathbb{C} ist algebraisch abgeschlossen.

Auch diesen Satz wollen wir hier nicht beweisen. Die Bezeichnung ‘‘Fundamentalsatz der Algebra’’ stammt aus den Anfangstagen der Algebra. Der erste strenge Beweis von Satz 16.15 stammt von Gauß aus dem Jahre 1799. Heute wird der Satz meist im Rahmen einer Funktionentheorievorlesung als Folgerung aus den Cauchy-schen Abschätzungsformeln bzw. aus dem Satz von Liouville hergeleitet (siehe z.B. Fischer/Lieb: Funktionentheorie). Dieser Beweis ist rein analytisch. Einen relativ elementaren mehr algebraischen Beweis findet man in dem Buch von Kostrikin [15]. Wir kommen in der Vorlesung *Einführung in die Algebra* darauf zurück.

Definition 16.16 . Es sei $P \in K[X]$ und $\lambda \in K$. Mit $v_\lambda(P) \in \mathbb{N}$ wird die größte natürliche Zahl k bezeichnet, so dass $P = (X - \lambda)^k Q$ für ein $Q \in K[X]$ gilt.

$v_\lambda(P)$ heißt die **Ordnung** von P in λ . Ist $k = v_\lambda(P) > 0$, so ist λ eine **Nullstelle der Ordnung** k von P oder eine **k -fache Nullstelle** von P .

Ist P **normiert**, d.h. der Leitkoeffizient von P gleich 1, und zerfällt P in Linearfaktoren, so gilt:

$$\begin{aligned} P &= \prod_{\lambda \in K} (X - \lambda)^{v_\lambda(P)} \\ &= \prod_{i=1}^m (X - \lambda_i)^{r_i}, \end{aligned}$$

wobei $\{\lambda_1, \dots, \lambda_m\} = \{\lambda \in K \mid v_\lambda(P) > 0\}$ und $r_i = v_{\lambda_i}(P)$.

Lemma 16.17 Es sei $F \in \text{End}(V)$ und $\lambda \in K$. Dann gilt:

$$v_\lambda(P_F) \geq \dim(\ker(F - \lambda \text{id}_V)).$$

Beweis: $V_0 = \ker(F - \lambda \text{id}_V)$ ist F -invariant und, F induziert auf V_0 die Abbildung λid_{V_0} . Man ergänze eine Basis (v_1, \dots, v_k) von $\ker(F - \lambda \text{id}_V)$ zu einer Basis \mathcal{B} von V . Dann hat die Matrix A von F bzgl. \mathcal{B} die Gestalt

$$A = \begin{pmatrix} \lambda E_k & B \\ 0 & A'' \end{pmatrix},$$

und somit ist

$$\begin{aligned} P_F = P_A &= \det \begin{pmatrix} (t - \lambda)E_k & -B \\ 0 & tE_{n-k} - A'' \end{pmatrix} \\ &= (t - \lambda)^k P_{A''}. \end{aligned}$$

Es folgt

$$\dim V_0 = k \leq v_\lambda(P_F).$$

□

Hieraus folgt nun sofort

Satz 16.18 Es sei $F \in \text{End}(V)$. Dann gilt: F ist genau dann diagonalisierbar, wenn gilt:

- a) P_F zerfällt über K in Linearfaktoren, und

b) \forall Eigenwerte $\lambda \in K$ von F gilt:

$$v_\lambda(P_F) = \dim \ker (F - \lambda id_V).$$

Beweis: Nach 16.4 und 16.7 ist F genau dann diagonalisierbar, wenn

$$V = \bigoplus_{\lambda \text{ Eigenwert von } F} \ker (F - \lambda id_V)$$

gilt. Mit Lemma 16.17 folgt jetzt leicht die Behauptung. \square

Korollar 16.19 Besitzt F n verschiedene Eigenwerte, so ist F diagonalisierbar.

Übungen

1. Seien $A, B \in M_n(K)$. Dann gilt: $P_{AB} = P_{BA}$.

Hinweis: Betrachte zunächst den Fall $A = \begin{pmatrix} E_r & 0 \\ 0 & 0 \end{pmatrix}$.

2. Sei $A \in M_n(K)$ invertierbar, und sei

$$P_A(t) = t^n + a_1 t^{n-1} + \dots + a_n.$$

Dann gilt:

$$P_{A^{-1}}(t) = \frac{1}{a_n} + \frac{a_1}{a_n} t + \dots + \frac{a_{n-1}}{a_n} t^{n-1} + t^n = (-1)^n \frac{t^n}{\det A} P_A \left(\frac{1}{t} \right)$$

und

$$P_{\tilde{A}}(t) = (-1)^n \frac{t^n}{\det A} P_A \left(\frac{\det A}{t} \right),$$

wobei \tilde{A} die zu A komplementäre Matrix ist.

3. Es sei K ein algebraisch abgeschlossener Körper (zum Beispiel $K = \mathbb{C}$). Es sei $A \in M_n(K)$ und $P_A = \prod_{i=1}^n (X - \lambda_i)$. Es sei $G \in K[X]$ und $B := G(A) \in K[A]$.

Zeigen Sie:

(a)

$$\det B = \prod_{i=1}^n G(\lambda_i).$$

Hinweis: G zerfällt in Linearfaktoren:

$$G = a \prod_{j=1}^m (X - \alpha_j), \quad a, \alpha_1, \dots, \alpha_m \in K.$$

Es folgt

$$B = a \prod_{j=1}^m (A - \alpha_j E_n).$$

Jetzt benutze man den Determinantenmultiplikationssatz und die Definition des charakteristischen Polynoms P_A .

(b)

$$P_B = \prod_{i=1}^n (X - G(\lambda_i)).$$

Hinweis: Für gegebenes $\lambda \in K$ betrachte man $G_\lambda = \lambda - G \in K[X]$. Auf $\lambda E - B = G_\lambda(A)$ wende man dann a) an.

(c)

$$\text{Spur } B = \sum_{i=1}^n G(\lambda_i).$$

4. (a) Es sei (i_1, \dots, i_n) eine Permutation von $(1, \dots, n)$ und $A = (e^{i_1}, \dots, e^{i_n}) \in M_n(\mathbb{C})$. Beweisen Sie: Es gibt ein $m \in \mathbb{N}$, $m > 0$, so dass $A^m = E_n$. Die Eigenwerte von A sind m -te Einheitswurzeln.

(b) Bestimmen Sie das charakteristische Polynom von

$$A = (e^n, e^1, e^2, \dots, e^{n-1}) \in M_n(\mathbb{C}).$$

Ist A über \mathbb{C} diagonalisierbar?

5. Es seien $\lambda_1, \dots, \lambda_n \in \mathbb{C}$ die Eigenwerte einer reellen $n \times n$ -Matrix $A \in M_n(\mathbb{R})$. Zeigen Sie:

$$\lambda_1^2 + \dots + \lambda_n^2 \in \mathbb{R}.$$

Hinweis: Benutzen Sie Aufgabe 3. c)!

6. Es sei V ein n -dimensionaler K -Vektorraum und $F \in \text{End}(V)$. Es gebe einen Vektor $v \in V$, so dass $(v, F(v), F^2(v), \dots, F^{n-1}(v))$ eine Basis von V ist. $\varphi: V \rightarrow K^n$ sei ein Isomorphismus, und es sei

$$a_0 = \varphi(v), a_1 = \varphi(F(v)), \dots, a_k = \varphi(F^k(v)), \dots$$

Dann gilt:

$$P_F = \frac{1}{\det(a_0, \dots, a_{n-1})} \det \begin{pmatrix} a_0 & a_1 & \dots & a_n \\ 1 & X & \dots & X^n \end{pmatrix}.$$

7. Es sei K ein Körper der Charakteristik p mit $p = 0$ oder $p > n$. Es sei $A \in M_n(K)$.

$$\tilde{B} = (XE_n - A) \in M_n(K[X])$$

sei die zu $(XE_n - A)$ komplementäre Matrix, und

$$P = \det(XE_n - A) \in K[X]$$

sei das charakteristische Polynom von A .

(a) Zeigen Sie:

$$\text{Spur } \tilde{B} = P',$$

wobei P' die Ableitung von P nach X bedeutet.

Hinweis: Zeigen Sie zunächst

$$\det(P^1, \dots, P^n)' = \sum_{i=1}^n \det(P^1, \dots, (P^i)', \dots, P^n)$$

für jede Matrix

$$M = (P^1, \dots, P^n) \in M_n(K[X])$$

mit den Spalten P^i . Wenden Sie dann dieses Ergebnis auf $P^i = Xe^i - a^i$ an, wobei a^i die i -te Spalte von A sei.

(b) Zeigen Sie: \tilde{B} hat die Gestalt

$$\tilde{B} = \tilde{B}_0 X^{n-1} + \tilde{B}_1 X^{n-2} + \dots + \tilde{B}_{n-1}$$

mit Matrizen $\tilde{B}_i \in M_n(K)$. Es sei

$$P = a_0 X^n + a_1 X^{n-1} + \dots + a_n \quad (a_0 = 1).$$

Aus der Gleichung

$$(XE_n - A)\tilde{B} = PE_n$$

und aus a) leite man die Formeln

$$\tilde{B}_0 = a_0 E_n, \quad \tilde{B}_i - A\tilde{B}_{i-1} = a_i E_n \quad \text{für } 1 \leq i \leq n-1, \quad -A\tilde{B}_{n-1} = a_n E_n$$

und

$$a_i = \frac{1}{n-i} \text{Spur}(\tilde{B}_i) \quad \text{für } i = 0, \dots, n-1$$

ab. Folgern Sie:

$$\tilde{A} = (-1)^{n-1} \tilde{B}_{n-1}$$

und

$$a_i = -\frac{1}{i} \text{Spur}(A\tilde{B}_{i-1}) \quad \text{für } i = 1, \dots, n.$$

17 Euklidische Vektorräume

In diesem Abschnitt bezeichnet V einen \mathbb{R} -Vektorraum.

Definition 17.1 Ein **Skalarprodukt** auf V ist eine Abbildung $g : V \times V \rightarrow \mathbb{R}$ mit folgenden Eigenschaften:

(a) g ist eine **Bilinearform**, d.h. die Abbildung g ist bilinear.

(b) g ist **symmetrisch**, d.h. für alle $v, w \in V$ gilt

$$g(v, w) = g(w, v).$$

(c) g ist **positiv definit**, d.h. für alle $v \in V$, $v \neq 0$ gilt

$$g(v, v) > 0.$$

Ein **euklidischer Vektorraum** ist ein endlich-dimensionaler \mathbb{R} -Vektorraum V zusammen mit einem Skalarprodukt g auf V .

$g(v, w)$ heißt das Skalarprodukt von v und w . An Stelle von $g(v, w)$ schreiben wir auch $\langle v, w \rangle$ oder $v \cdot w$.

Beispiel 17.2 a) \mathbb{R}^n mit dem Standardskalarprodukt

$$x \cdot y = \langle x, y \rangle = \sum_{j=1}^n x_j y_j$$

ist ein euklidischer Vektorraum.

b) Es sei $g_n(x) = \cos(nx)$, $f_n(x) = \sin(nx)$. Wir betrachten den von den Funktionen g_0, g_1, g_2, \dots und f_1, f_2, \dots erzeugten Untervektorraum \mathcal{V} von $\text{Abb}([0, 2\pi], \mathbb{R})$. Die Elemente von \mathcal{V} werden auch trigonometrische Polynome oder Fourier-Polynome genannt. Für $f, g \in \mathcal{V}$ sei

$$\langle f, g \rangle := \int_0^{2\pi} f(x)g(x) dx.$$

Dies ist ein Skalarprodukt auf \mathcal{V} , denn

$$\langle f, f \rangle = \int_0^{2\pi} f(x)^2 dx > 0, \text{ wenn } f \neq 0.$$

Aus der Integralrechnung sind die Integrale $\langle f_n, f_m \rangle$, $\langle f_n, g_m \rangle$, $\langle g_n, g_m \rangle$ bekannt.

So ist zum Beispiel für $n > 0$:

$$\begin{aligned}\langle f_n, f_n \rangle &= \int_0^{2\pi} \sin^2(nx) dx \stackrel{[y=nx]}{=} \frac{1}{n} \int_0^{2\pi n} \sin^2 y dy \\ &= \int_0^{2\pi} \sin^2 y dy = - \int_0^{2\pi} \sin y d(\cos y) \\ &\stackrel{[\text{part.Int.}]}{=} - \sin y \cos y \Big|_0^{2\pi} + \int_0^{2\pi} \cos^2 y dy \\ &= \int_0^{2\pi} \cos^2 y dy = \langle g_n, g_n \rangle\end{aligned}$$

Da $\cos^2 + \sin^2 = 1$, folgt

$$\begin{aligned}2\langle f_n, f_n \rangle &= \langle f_n, f_n \rangle + \langle g_n, g_n \rangle = \int_0^{2\pi} (\sin^2 y + \cos^2 y) dy \\ &= \int_0^{2\pi} dy = 2\pi\end{aligned}$$

und somit $\langle f_n, f_n \rangle = \langle g_n, g_n \rangle = \pi$ für $n > 0$.

Weiter kann man zeigen:

Für $n \neq m$ ist $\langle f_n, f_m \rangle = \langle g_n, g_m \rangle = 0$ und es gilt auch $\langle f_n, g_m \rangle = 0$ für alle $n > 0, m \geq 0$.

Ist nun $f \in \mathcal{V}$ ein beliebiges Element, so kann man f in der Form

$$f(x) = a_0 + \sum_{n=1}^N (a_n \cos(nx) + b_n \sin(nx))$$

schreiben, also

$$f = a_0 g_0 + \sum_{n=1}^N a_n g_n + \sum_{n=1}^N b_n f_n.$$

Man kann die Koeffizienten a_i, b_i als Skalarprodukte berechnen:

$$\begin{aligned}\langle f, g_0 \rangle &= a_0 \langle g_0, g_0 \rangle + \sum_{n=1}^N a_n \langle g_n, g_0 \rangle + \sum_{n=1}^N b_n \langle f_n, g_0 \rangle \\ &= a_0 \cdot 2\pi\end{aligned}$$

also $a_0 = \frac{1}{2\pi} \langle f, g_0 \rangle$. Genauso folgt $a_n = \frac{1}{\pi} \langle f, g_n \rangle, b_n = \frac{1}{\pi} \langle f, f_n \rangle$.

Insbesondere folgt, dass die Funktionen $g_0, g_1, \dots, f_1, f_2, \dots$ eine Basis des Vektorraums \mathcal{V} bilden. Der Grund hierfür war die Eigenschaft: $\langle f_n, f_m \rangle = \langle g_n, g_m \rangle = 0$ für $n \neq m$ und $\langle g_n, f_m \rangle = 0$ für alle n, m .

- c) Es sei $\Omega = \{x_1, \dots, x_n\}$ eine endliche Menge und $p : \Omega \rightarrow \mathbb{R}_+$ eine Funktion mit

$$\sum_{x \in \Omega} p(x) = 1.$$

Wir interpretieren $p_i = p(x_i)$ als die **Wahrscheinlichkeit** für das Ereignis x_i . Eine Funktion $f : \Omega \rightarrow \mathbb{R}$ heißt **reelle Zufallsvariable** auf Ω . Die Zahl

$$E(f) := \sum_{x \in \Omega} p(x) f(x) \in \mathbb{R}$$

heißt der **Erwartungswert** von f .

Die Zufallsvariablen bilden den n -dimensionalen Vektorraum $\mathcal{V} = \text{Abb}(\Omega, \mathbb{R})$

$$\langle f, g \rangle := \sum_{x \in \Omega} p(x) f(x) g(x) = E(fg)$$

ist ein Skalarprodukt auf \mathcal{V} , da $p_i > 0$ für alle $i = 1, \dots, n$.

Definition 17.3 Es sei (V, g) n -dimensionaler euklidischer Vektorraum. (v_1, \dots, v_n) sei eine Basis von V . Dann heißt

$$G = G(v_1, \dots, v_n) = (g(v_i, v_j))_{i,j=1,\dots,n} \in M_n(\mathbb{R})$$

die **Gramsche Matrix** von v_1, \dots, v_n bezüglich g oder die **Matrix** von g in der Basis (v_1, \dots, v_n) . $\det G$ heißt die **Gramsche Determinante** von v_1, \dots, v_n bezüglich g .

Man kann $g(v, w)$ für alle $v, w \in V$ mit Hilfe der Matrix G berechnen.

Sei dazu $x = [x_1, \dots, x_n]$ der Koordinatenvektor von v bezüglich der Basis $\mathcal{B} = (v_1, \dots, v_n)$, also

$$v = \sum_{i=1}^n x_i v_i$$

und $y = [y_1, \dots, y_n]$ sei der Koordinatenvektor von w ,

$$w = \sum_{i=1}^n y_i v_i.$$

Aus der Bilinearität von g folgt dann

$$g(v, w) = \sum_{i,j=1}^n x_i y_j g(v_i, v_j) = x^t G y.$$

Da g symmetrisch ist, gilt insbesondere $g(v_i, v_j) = g(v_j, v_i)$; G ist also eine symmetrische Matrix.

Da g positiv definit ist, ist auch die Matrix G positiv definit, d.h. es gilt

$$x^t G x > 0 \text{ für alle } x \in \mathbb{R}^n \setminus \{0\}.$$

Wir haben damit den folgenden

Satz 17.4 Sei V n -dimensionaler \mathbb{R} -Vektorraum und $\mathcal{B} = (v_1, \dots, v_n)$ sei eine Basis von V .

Dann sind

$$S_2(V) = \{g : V \times V \longrightarrow \mathbb{R} \mid g \text{ symmetrisch und bilinear}\}$$

und

$$S_n(\mathbb{R}) = \{A \in M_n(\mathbb{R}) \mid A = A^t\}$$

\mathbb{R} -Vektorräume (Unterräume von $\text{Abb}(V \times V, \mathbb{R})$ beziehungsweise von $M_n(\mathbb{R})$).

Die Abbildung

$$g \longmapsto G = (g(v_i, v_j))_{i,j=1,\dots,n}$$

ist ein Isomorphismus

$$S_2(V) \longrightarrow S_n(\mathbb{R}).$$

Die Umkehrabbildung ist durch die Formel

$$g(v, w) = x^t G y \text{ für } x = \sum_{i=1}^n x_i v_i, \quad y = \sum_{i=1}^n y_i v_i$$

gegeben. g ist genau dann positiv definit, wenn G es ist. □

Wir untersuchen das Transformationsverhalten von Bilinearformen $g : V \times V \longrightarrow \mathbb{R}$ unter Basiswechsel.

Zunächst bemerken wir, dass für Endomorphismen $f : V \rightarrow V$ gilt: Sind A, B Matrizen von f bezüglich verschiedener Basen $\mathcal{B} = (v_1, \dots, v_n), \mathcal{B}' = (v'_1, \dots, v'_n)$ und ist S die Basiswechselmatrix, $v'_j = \sum_{i=1}^n s_{ij} v_i$, so gilt

$$B = S^{-1} A S.$$

Anders verhält es sich bei Bilinearformen:

Satz 17.5 (Transformationsformel für Bilinearformen)

Es seien $\mathcal{B} = (v_1, \dots, v_n), \mathcal{B}' = (v'_1, \dots, v'_n)$ Basen von V , $g : V \times V \rightarrow \mathbb{R}$ sei eine Bilinearform. Es sei $G = (g(v_i, v_j))_{i,j=1,\dots,n}$ und $G' = (g(v'_i, v'_j))_{i,j=1,\dots,n}$. Weiter sei $S = (s_{ij})$ die Basiswechselmatrix mit $v'_j = \sum_{i=1}^n s_{ij} v_i$. Dann gilt

$$G' = S^t G S.$$

Beweis: Der Beweis ist reine Routine:

$$\begin{aligned} v &= \sum x_i v_i, & w &= \sum y_i v_i \text{ und} \\ v &= \sum x'_j v'_j = \sum s_{ij} x'_j v_i, \\ w &= \sum y'_j v'_j = \sum s_{ij} y'_j v_i \end{aligned}$$

impliziert $x = Sx'$, $y = Sy'$ und somit

$$x^t G' y' = g(v, w) = x^t G y = (Sx')^t G S y' = x'^t S^t G S y'.$$

Da dies für alle $x', y' \in \mathbb{R}^n$ gilt, folgt

$$G' = S^t G S.$$

□

Definition 17.6 Es sei V ein euklidischer Vektorraum mit Skalarprodukt $\langle \cdot, \cdot \rangle$. Für $v \in V$ heißt $\|v\| := \sqrt{\langle v, v \rangle}$ die (**euklidische**) **Norm** von v .

Wichtig ist nun der folgende

Satz 17.7 (Cauchy-Schwarzsche Ungleichung)

Für alle $v, w \in V$ gilt

$$|\langle v, w \rangle| \leq \|v\| \|w\|. \quad (50)$$

Gleichheit gilt genau dann, wenn v, w linear abhängig sind.

Beweis:

1. Ist $w = 0$, so gilt in (50) Gleichheit und v und w sind linear abhängig.
2. Es sei $w \neq 0$. Für alle $t \in \mathbb{R}$ gilt

$$0 \leq \langle v + tw, v + tw \rangle = \|v\|^2 + 2\langle v, w \rangle t + t^2 \|w\|^2.$$

Das quadratische Polynom

$$f(t) = t^2 + \frac{2\langle v, w \rangle}{\|w\|^2} t + \frac{\|v\|^2}{\|w\|^2}$$

besitzt also höchstens eine reelle Nullstelle, die Diskriminante

$$\Delta = \frac{\langle v, w \rangle^2 - \|v\|^2 \|w\|^2}{\|w\|^4}$$

von f ist daher nicht positiv, es gilt also

$$\langle v, w \rangle^2 \leq \|v\|^2 \|w\|^2,$$

d.h. $|\langle v, w \rangle| \leq \|v\| \|w\|$.

Ist $\Delta = 0$, gilt also $|\langle v, w \rangle| = \|v\| \|w\|$, so gibt es ein $t \in \mathbb{R}$ mit $f(t) = 0$, also

$$\langle v + tw, v + tw \rangle = 0$$

und somit $v + tw = 0$. v, w sind dann linear abhängig. □

Korollar 17.8 $\| \cdot \|$ ist eine **Norm** auf V , d.h. es gilt

- (a) $\|\lambda v\| = |\lambda| \|v\|$ für alle $\lambda \in \mathbb{R}, v \in V$
 (b) $\|v + w\| \leq \|v\| + \|w\|$ für alle $v, w \in V$
 (c) $\|v\| \geq 0$ und $\|v\| = 0$ nur wenn $v = 0$.

Beweis: Lediglich die Dreiecksungleichung (b) bedarf eines Beweises:

$$\begin{aligned} \|v + w\|^2 &= \|v\|^2 + 2\langle v, w \rangle + \|w\|^2 \\ &\leq \|v\|^2 + 2\|v\| \|w\| + \|w\|^2 \\ &= (\|v\| + \|w\|)^2 \end{aligned}$$

□

Definition 17.9 Es seien $v, w \in V \setminus \{0\}$. Dann gibt es genau eine reelle Zahl $\alpha \in [0, \pi]$, so dass

$$\cos \alpha = \frac{\langle v, w \rangle}{\|v\| \|w\|}.$$

Diese Zahl α heißt der (unorientierte) **Winkel** zwischen v und w , in Zeichen:

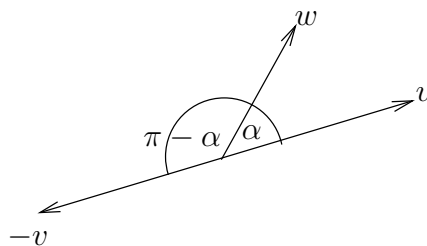
$$\alpha = \sphericalangle(v, w).$$

Es gilt

$$\begin{aligned} \alpha = 0 &\iff v = \lambda w \text{ mit } \lambda > 0 \\ \alpha = \pi &\iff v = \lambda w \text{ mit } \lambda < 0 \\ \alpha = \frac{\pi}{2} &\iff \langle v, w \rangle = 0 \\ 0 < \alpha < \frac{\pi}{2} &\iff \langle v, w \rangle > 0, \quad v, w \text{ linear unabhängig} \\ &\hspace{15em} \text{(spitzer Winkel)} \\ \frac{\pi}{2} < \alpha < \pi &\iff \langle v, w \rangle < 0, \quad v, w \text{ linear unabhängig} \\ &\hspace{15em} \text{(stumpfer Winkel)} \end{aligned}$$

Da $\cos(\pi - \alpha) = -\cos \alpha$ gilt ergibt sich

$$\sphericalangle(-v, w) = \pi - \sphericalangle(v, w) \quad \text{(komplementäre Winkel)}$$



Beispiel 17.10 (Kosinussatz)

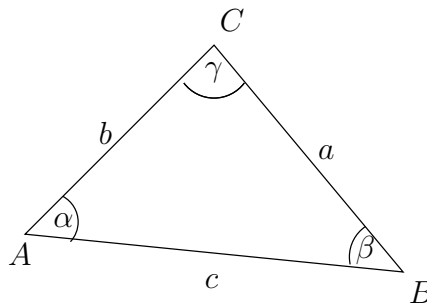
Ist $\alpha = \sphericalangle(v, w)$, so gilt

$$\|v - w\|^2 = \|v\|^2 + \|w\|^2 - 2\|v\| \|w\| \cos \alpha \quad (51)$$

Dies folgt sofort aus der Definition von α und der Norm:

$$\begin{aligned} \|v - w\|^2 &= \langle v - w, v - w \rangle = \langle v, v \rangle - 2\langle v, w \rangle + \langle w, w \rangle \\ &= \|v\|^2 - 2\|v\| \|w\| \cos \alpha + \|w\|^2. \end{aligned}$$

Dieser Satz ist aus der Dreiecksgeometrie bekannt: Es sei ein Dreieck mit den Ecken A, B, C gegeben. Die Seiten a, b, c und die Winkel α, β, γ seien wie im Bild gewählt.



Man bekommt mit $v = \vec{AB}, w = \vec{AC}$:

$$v - w = \vec{AB} - \vec{AC} = \vec{CA} + \vec{AB} = \vec{CB}$$

und $\|v\| = c, \|w\| = b, \|v - w\| = a$.

Damit folgt aus (51):

$$a^2 = b^2 + c^2 - 2bc \cos \alpha.$$

Analog gilt natürlich auch

$$\begin{aligned} b^2 &= a^2 + c^2 - 2ac \cos \beta \\ c^2 &= a^2 + b^2 - 2ab \cos \gamma. \end{aligned}$$

Hieraus folgt auch der Satz von Pythagoras. Ist $\gamma = \frac{\pi}{2}$, das Dreieck also rechtwinklig, so gilt $\cos \gamma = 0$ und damit

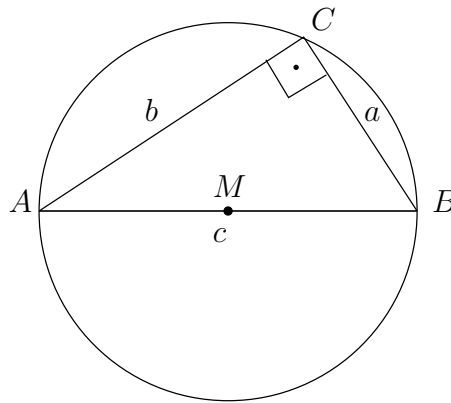
$$c^2 = a^2 + b^2.$$

Beispiel 17.11 (Satz von Thales)

Sind v, w linear unabhängig mit $\|v\| = \|w\|$, so gilt $v - w$ ist orthogonal zu $v + w$, d.h. $\sphericalangle(v - w, v + w) = \frac{\pi}{2}$. Das ist klar: Es gilt ja

$$\langle v - w, v + w \rangle = \langle v, v \rangle - \langle w, w \rangle = 0.$$

Ist ABC ein Dreieck auf einem Kreis, so dass die Seite $c = AB$ ein Durchmesser ist, so ist ABC rechtwinklig, $\gamma = \frac{\pi}{2}$.



Man setze $v = \vec{MB}$, $w = \vec{MC}$; dann ist $-v = \vec{MA}$, $v - w = \vec{CB}$ und $v + w = \vec{AM} + \vec{MC} = \vec{AC}$. Es gilt also \vec{CB} ist senkrecht zu \vec{AC} , weil

$$\|\vec{MC}\| = \|w\| = \|v\| = \|\vec{MB}\|.$$

Es sei jetzt V ein n -dimensionaler euklidischer Vektorraum.

Definition 17.12 a) $v, w \in V$ heißen **orthogonal**, wenn $\langle v, w \rangle = 0$ gilt.

b) Zwei Untervektorräume U, W von V heißen **orthogonal**, wenn $\langle v, w \rangle = 0$ für alle $v \in U$, $w \in W$ gilt.

c) Ist $U \subset V$ ein Untervektorraum, so heißt

$$U^\perp = \{v \in V \mid \langle v, w \rangle = 0 \text{ für alle } w \in U\}$$

das **orthogonale Komplement** von U in V .

d) Ein System (v_1, \dots, v_m) von Vektoren in V heißt **orthonormal**, wenn $\langle v_i, v_j \rangle = \delta_{ij}$ für alle $i, j = 1, \dots, m$ gilt, d.h. wenn die Vektoren paarweise zueinander orthogonal sind und wenn $\|v_i\| = 1$ für $i = 1, \dots, m$. Eine **Orthonormalbasis** ist eine Basis, die orthonormal ist.

Satz 17.13 Es sei $U \subset V$ ein Untervektorraum. Dann gilt:

a) U^\perp ist Untervektorraum von U und es gilt $V = U \oplus U^\perp$.

b) $(U^\perp)^\perp = U$

Beweis: Zu a) Offensichtlich ist U^\perp ein Untervektorraum von U , denn das Skalarprodukt ist bilinear.

Ist nun $v \in U \cap U^\perp$, so ist insbesondere $\langle v, v \rangle = 0$, also $v = 0$. Damit ist $U \cap U^\perp = 0$. Damit ist die Summe von U und U^\perp direkt. Um $V = U \oplus U^\perp$ zu zeigen, genügt es $\dim U^\perp = n - \dim U$ zu beweisen.

Es sei dazu (u_1, \dots, u_m) eine Basis von U . Dann gilt offensichtlich $v \in U^\perp \iff \langle v, u_j \rangle = 0$ für $j = 1, \dots, m$.

Wir ergänzen (u_1, \dots, u_m) zu einer Basis (u_1, \dots, u_n) von V und bezeichnen mit $G = (g_{ij})$, $g_{ij} = \langle u_i, u_j \rangle$ die Gramsche Matrix. Ist nun $v = \sum_{i=1}^n x_i u_i$, $x = [x_1, \dots, x_n]$, so gilt also

$$v \in U^\perp \iff \sum_{i=1}^n g_{ij} x_i = 0 \text{ für } j = 1, \dots, m.$$

Nun hat aber die Matrix G den Rang n , denn für $v = \sum_{i=1}^n x_i u_i$ gilt: Ist $Gx = 0$, so ist auch $\langle v, v \rangle = x^t Gx = 0$, also $v = 0$ und somit $x = 0$. Folglich hat G den Rang n und $(g_{ij})_{\substack{i=1, \dots, n \\ j=1, \dots, m}}$ den Rang m .

U^\perp , die Lösungsmenge des linearen Gleichungssystems

$$\sum_{i=1}^n g_{ij} x_i = 0 \text{ für } j = 1, \dots, m,$$

hat also die Dimension $n - m$.

$V = U \oplus U^\perp$ ist bewiesen.

Zu b) Es gilt $\dim(U^\perp)^\perp = n - \dim U^\perp = \dim U$ und $U \subset (U^\perp)^\perp$, folglich $U = (U^\perp)^\perp$.
□

Beispiel 17.14 Es sei \mathbb{R}^3 mit dem kanonischen Skalarprodukt versehen.

a) Es sei $U = \{(x, y, z) \in \mathbb{R}^3 \mid 3x + 2y - z = 0\}$. Wie berechnet man U^\perp ?

Es sei $v := (3, 2, -1)$. Dann gilt

$$\begin{aligned} u = (x, y, z) \in U &\iff \langle u, v \rangle = 0 \\ &\iff u \in (\mathbb{R}v)^\perp \end{aligned}$$

Also ist $U = (\mathbb{R}v)^\perp$ und somit $U^\perp = (\mathbb{R}v)^{\perp\perp} = \mathbb{R}v$.

b) Es sei U der von $u_1 = (1, 1, 0)$, $u_2 = (1, 0, 2)$ aufgespannte Untervektorraum von \mathbb{R}^3 .

Dann gilt

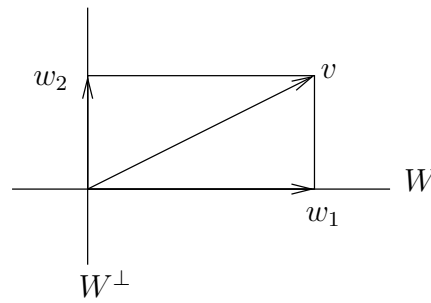
$$v = (x, y, z) \in U^\perp \iff \begin{cases} x + y = 0 \\ x + 2z = 0 \end{cases}$$

Eine Lösung ist $(1, -1, -\frac{1}{2})$ und somit gilt $U^\perp = \mathbb{R}(1, -1, -\frac{1}{2})$.

Definition 17.15 Es sei W ein Untervektorraum von V . Für jeden Vektor $v \in V$ gibt es eine eindeutige Zerlegung $v = w_1 + w_2$ mit $w_1 \in W$ und $w_2 \in W^\perp$.

w_1 heißt dann die **orthogonale Projektion** von v auf W .

w_2 ist dann die orthogonale Projektion von v auf W^\perp .



Die Abbildung $p_W \in \text{End}(V)$ mit $p_W(v) = w_1$ heißt der **orthogonale Projektor** von V auf W .

Es gilt

$$\text{im}(p_W) = W, \quad p_W^2 = p_W,$$

d.h. p_W ist idempotent. Man hat offensichtlich die Zerlegung

$$\text{id}_V = p_W + p_{W^\perp}$$

der Identität. Weiter gilt:

$$p_W \circ p_{W^\perp} = p_{W^\perp} \circ p_W = 0.$$

Lemma 17.16 Jedes Orthonormalsystem ist linear unabhängig.

Beweis: Es sei (v_1, \dots, v_m) ein Orthonormalsystem. Ist nun $\sum_{i=1}^m a_i v_i = 0$, so ist auch

$$0 = \left(\sum_{i=1}^m a_i v_i, v_j \right) = \sum_{i=1}^m a_i (v_i, v_j) = \sum_{i=1}^m a_i \delta_{ij} = a_j$$

für alle $j = 1, \dots, m$. □

Satz 17.17 Jeder euklidische Vektorraum besitzt eine Orthonormalbasis.

Beweis: Induktion nach $n = \dim V$.

1. Ist $n = 1$ und $v \in V \setminus \{0\}$, so ist v_1 mit $v_1 := \frac{1}{\|v\|}v$ eine Orthonormalbasis.
2. $n - 1 \rightarrow n$: Es sei $n = \dim V$, $n \geq 2$.

Wähle einen Vektor $v_1 \in V$ mit $\|v_1\| = 1$. Dann ist $W := (\mathbb{R}v_1)^\perp$ $(n-1)$ -dimensional. Nach Induktionsvoraussetzung hat W eine Orthonormalbasis v_2, \dots, v_n .

Offensichtlich ist dann v_1, v_2, \dots, v_n eine Orthonormalbasis von V . □

Definition 17.18 Es sei $\mathcal{B} = (v_1, \dots, v_n)$ eine Orthonormalbasis von V . Für $v \in V$ heißt dann $\langle v, v_i \rangle$ der i -te **Fourierkoeffizient** von v (bzgl. \mathcal{B}).

Es gilt

$$v = \langle v, v_1 \rangle v_1 + \dots + \langle v, v_n \rangle v_n.$$

Die Abbildung $\Phi : V \rightarrow \mathbb{R}^n$ mit

$$\Phi(v) = (x_1, \dots, x_n), \text{ wenn } x_i = \langle v, v_i \rangle$$

ist ein \mathbb{R} -Vektorraumisomorphismus und es gilt

$$\langle v, w \rangle = \langle \Phi(v), \Phi(w) \rangle.$$

Das Skalarprodukt von v und w ist ja

$$\begin{aligned} \langle v, w \rangle &= \left\langle \sum_i \langle v, v_i \rangle v_i, \sum_j \langle w, v_j \rangle v_j \right\rangle \\ &= \sum_{i,j} \langle v, v_i \rangle \langle w, v_j \rangle \delta_{ij} = \sum_i \langle v, v_i \rangle \langle w, v_i \rangle. \end{aligned}$$

Der euklidische Raum V ist isomorph zu \mathbb{R}^n mit dem Standardskalarprodukt.

Wir sehen also: Bis auf Isomorphie gibt es nur einen n -dimensionalen euklidischen Raum.

Lemma 17.19 Es sei W ein k -dimensionaler Untervektorraum von V und (v_1, \dots, v_k) sei eine Orthonormalbasis von W . Dann kann man die orthogonale Projektion

$$p_W : V \rightarrow V$$

wie folgt beschreiben:

$$p_W(v) = \sum_{i=1}^k \langle v, v_i \rangle v_i$$

Beweis: Es gilt ja für $j = 1, \dots, k$:

$$\left\langle v - \sum_{i=1}^k \langle v, v_i \rangle v_i, v_j \right\rangle = 0.$$

□

Wir erklären jetzt eine Methode, wie man aus einer gegebenen Basis von V eine Orthonormalbasis von V gewinnt.

Satz 17.20 (Schmidtsches Orthonormalisierungsverfahren)

Es sei (w_1, \dots, w_n) eine Basis von V .

Man setze

$$v_1 := \frac{1}{\|w_1\|} w_1,$$

und für $k \geq 1$:

$$v_{k+1} := \frac{w_{k+1} - \sum_{i=1}^k \langle w_{k+1}, v_i \rangle v_i}{\|w_{k+1} - \sum_{i=1}^k \langle w_{k+1}, v_i \rangle v_i\|}$$

Dann ist (v_1, \dots, v_n) eine Orthonormalbasis von V mit

$\text{Span}(w_1, \dots, w_k) = \text{Span}(v_1, \dots, v_k)$ für $k = 1, \dots, n$.

Beweis: Es sei $V_k := \text{Span}(w_1, \dots, w_k)$. Durch Induktion nach k zeigen wir, dass (v_1, \dots, v_k) eine Orthonormalbasis von V_k ist.

- a) $k = 1$: Es gilt: $V_1 = \mathbb{R}w_1 = \mathbb{R}v_1$ und $\|v_1\| = 1$.
- b) Es sei $1 \leq k \leq n - 1$ und es sei schon gezeigt, dass (v_1, \dots, v_k) eine Orthonormalbasis von V_k ist. Nach Lemma 17.19 ist

$$\tilde{w}_{k+1} := \sum_{i=1}^k \langle w_{k+1}, v_i \rangle v_i \in V_k$$

die orthogonale Projektion von w_{k+1} auf V_k . Da $w_{k+1} \notin V_k$, ist $w_{k+1} - \tilde{w}_{k+1} \in V_k^\perp \setminus \{0\}$.

Also ist auch

$$v_{k+1} := \frac{w_{k+1} - \tilde{w}_{k+1}}{\|w_{k+1} - \tilde{w}_{k+1}\|} \in V_k^\perp \text{ mit } \|v_{k+1}\| = 1.$$

Da außerdem $v_{k+1} \in V_{k+1}$ (weil $w_{k+1}, \tilde{w}_{k+1} \in V_{k+1}$), ist (v_1, \dots, v_{k+1}) eine Orthonormalbasis von V_{k+1} . \square

Beispiel 17.21 Um aus einer beliebigen Basis (w_1, \dots, w_n) von V eine Orthogonalbasis (v_1, \dots, v_n) zu finden, derart, dass

$$\text{Span}(v_1, \dots, v_k) = V_k,$$

wobei $V_k := \text{Span}(w_1, \dots, w_k)$ ist, kann man auch so vorgehen. Man setzt $v_1 = w_1$. Für $k = 1, \dots, n - 1$ bestimmt man $x = (x_1, \dots, x_k)$, so dass

$$v_{k+1} = w_{k+1} - \sum_{i=1}^k x_i w_i$$

orthogonal zu V_k ist; das liefert die linearen Bedingungen

$$\langle v_{k+1}, w_j \rangle = 0 \text{ für } j = 1, \dots, k$$

und somit das inhomogene lineare Gleichungssystem

$$\sum_{i=1}^k x_i \langle w_i, w_j \rangle = \langle w_{k+1}, w_j \rangle, \quad j = 1, \dots, k \quad (52)$$

mit k Gleichungen in den Unbekannten x_1, \dots, x_k . Die Koeffizienten des Gleichungssystems entnimmt man der Gramschen Matrix

$$G = (g_{ij}), \quad g_{ij} = \langle w_i, w_j \rangle$$

von (w_1, \dots, w_n) .

Das System (52) hat die Form

$$G_k x = b_k, \quad (53)$$

wobei $G_k = (g_{ij})_{i,j \leq k}$ Untermatrix von G und $b_k = [g_{k+1,1}, \dots, g_{k+1,k}]$ ist.

Da G_k invertierbar ist, gibt es genau eine Lösung x von (53).

Die so gewonnene Basis ist orthogonal. Man kann sie normalisieren, in dem man v_i durch $\frac{1}{\|v_i\|}v_i$ ersetzt.

Ein konkretes Beispiel:

Es sei V der Vektorraum der Polynome $a_0 + a_1x + a_2x^2 + a_3x^3$ vom Grad ≤ 3 und das Skalarprodukt auf V sei

$$\langle f, g \rangle = \int_{-1}^1 f(x)g(x) dx.$$

Wir wollen die Basis $\mathcal{B} = (1, x, x^2, x^3)$ orthogonalisieren:

Die Gramsche Matrix G von \mathcal{B} ist

$$\begin{aligned} G &= (g_{ij})_{i,j=1,\dots,4} \text{ mit } g_{ij} = \int_{-1}^1 x^{i+j-2} dx = \left[\frac{1}{i+j-1} x^{i+j-1} \right]_{-1}^1 \\ &= \frac{1}{i+j-1} (1 + (-1)^{i+j}), \end{aligned}$$

also

$$G = \begin{pmatrix} 2 & 0 & \frac{2}{3} & 0 \\ 0 & \frac{2}{3} & 0 & \frac{2}{5} \\ \frac{2}{3} & 0 & \frac{2}{5} & 0 \\ 0 & \frac{2}{5} & 0 & \frac{2}{7} \end{pmatrix}.$$

Das oben beschriebene Verfahren liefert:

$$v_1 = 1.$$

$$v_2 = x - x_1 \cdot 1 \text{ und } g_{11}x_1 = g_{21}, \text{ also } 2x_1 = 0, \text{ und somit } x_1 = 0, \\ \text{d.h. } v_2 = x.$$

$$v_3 = x^2 - x_1 - x_2x$$

und

$$\begin{pmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} \frac{2}{3} \\ 0 \end{pmatrix},$$

d.h.

$$\begin{pmatrix} 2 & 0 \\ 0 & \frac{2}{3} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} \frac{2}{3} \\ 0 \end{pmatrix},$$

$$\text{also } x_1 = \frac{1}{3}, x_2 = 0 \text{ und } v_3 = x^2 - \frac{1}{3}.$$

Schließlich

$$v_4 = x^3 - x_1 - x_2x - x_3x^2$$

und

$$\begin{pmatrix} 2 & 0 & \frac{2}{3} \\ 0 & \frac{2}{3} & 0 \\ \frac{2}{3} & 0 & \frac{2}{5} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ \frac{2}{5} \\ 0 \end{pmatrix}.$$

Mit Gauß-Elimination ergibt sich

$$\begin{pmatrix} 2 & 0 & \frac{2}{3} \\ 0 & \frac{2}{3} & 0 \\ 0 & 0 & \frac{2}{5} - \frac{2}{9} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ \frac{2}{5} \\ 0 \end{pmatrix}.$$

Also ist $x_3 = 0$, $x_2 = \frac{3}{5}$, $x_3 = 0$ und somit $v_4 = x^3 - \frac{3}{5}x$.

Damit ist

$$(v_1, v_2, v_3, v_4) = \left(1, x, x^2 - \frac{1}{3}, x^3 - \frac{3}{5}x\right)$$

ein orthogonale Basis von V . Allerdings ist diese Basis nicht normiert.

Das folgende Kriterium von Hurwitz ist nützlich, um eine symmetrische Matrix auf positive Definitheit zu prüfen.

Satz 17.22 Eine symmetrische Matrix $G = (g_{ij}) \in M_n(\mathbb{R})$ ist genau dann positiv definit, wenn die Hauptminoren $\det G_k$, $k = 1, \dots, n$ positiv sind.

Dabei ist $G_k = (g_{ij})_{i,j \leq k} \in M_k(\mathbb{R})$ die k -te Hauptuntermatrix von G .

Beweis: Es sei $g : V \times V \rightarrow \mathbb{R}$ eine symmetrische Bilinearform, (w_1, \dots, w_n) eine Basis von V und G sei die Gramsche Matrix von (w_1, \dots, w_n) bezüglich g . Es sei weiter V_k der von w_1, \dots, w_k aufgespannte Unterraum von V und g_k die Einschränkung von g auf V_k . Dann ist G_k die Gramsche Matrix von (w_1, \dots, w_k) bezüglich g_k .

1. Es sei g positiv definit. Dann sind auch die Einschränkungen g_k positiv definit. Nach dem Orthonormalisierungsverfahren gibt es eine invertierbare Matrix S_k , so dass $S_k^t G_k S_k$ die Einheitsmatrix ist. Es gilt also $1 = \det(S_k^t G_k S_k) = (\det S_k)^2 \det G_k$ und folglich ist $\det G_k > 0$.
2. Es gelte jetzt $\det G_k > 0$ für $k = 1, \dots, n$. Induktiv zeigen wir, dass G_k positiv definit ist für $k = 1, \dots, n$.

(a) $k = 1$: $G_1 = (g_{11})$, $g_{11} > 0 \implies G_1$ positiv definit

(b) $k \rightarrow k + 1$: Es sei G_k positiv definit, also g_k positiv definit. Wähle eine Orthonormalbasis (v_1, \dots, v_k) von V_k . Da G_k nicht singulär ist, denn $\det G_k \neq 0$, kann man nach 17.21 genau ein $v_{k+1} \in V_{k+1}$ finden, so dass

$$v_{k+1} = w_{k+1} - \sum_{i=1}^k x_i w_i$$

und

$$\sum_{i=1}^k g_{ij} x_i = g_{k+1,j} \text{ für } j = 1, \dots, k.$$

Die Gramsche Matrix von (v_1, \dots, v_{k+1}) bezüglich g_{k+1} hat dann die Form

$$G' = \begin{pmatrix} 1 & \dots & 0 \\ & \ddots & \vdots \\ \vdots & & 1 & 0 \\ 0 & \dots & 0 & \gamma \end{pmatrix} = S^t G_{k+1} S.$$

Es gilt dann

$$\gamma = \det G' = (\det S)^2 \det G_{k+1} > 0.$$

Damit ist gezeigt, dass G' positiv definit ist und somit natürlich auch G_{k+1} . \square

Bemerkung. Der Beweis hat eigentlich mehr gezeigt: Ist $\det G_k \neq 0$ für $k = 1, \dots, n$, so gibt es eine Basis (v_1, \dots, v_n) von V , so dass die Matrix $(g(v_i, v_j))_{i,j}$ Diagonalmatrix der Form

$$\left(\begin{array}{ccccccc} 1 & & & & & & \\ & \ddots & & & & & \\ & & 1 & & & & \\ & & & -1 & & & \\ & & & & \ddots & & \\ & & & & & & -1 \end{array} \right) \left. \begin{array}{l} \left. \vphantom{\begin{pmatrix} 1 \\ \vdots \\ 1 \\ -1 \\ \vdots \\ -1 \end{pmatrix}} \right\} r \\ \left. \vphantom{\begin{pmatrix} 1 \\ \vdots \\ 1 \\ -1 \\ \vdots \\ -1 \end{pmatrix}} \right\} s \end{array} \right\}$$

ist, wobei r die Anzahl der positiven und s die Anzahl der negativen Elemente in der Folge

$$\left(\det G_1, \frac{\det G_2}{\det G_1}, \dots, \frac{\det G_n}{\det G_{n-1}} \right)$$

ist. Wir kommen darauf in Abschnitt 22 über quadratische Formen zurück.

Übungen

1. Welche der folgenden Abbildungen definiert ein Skalarprodukt auf \mathbb{R}^3 ?

(a) $\langle x, y \rangle_A := x^t A y$ mit

$$A = \begin{pmatrix} 1 & 2 & 0 \\ 2 & 5 & 1 \\ 0 & 1 & 2 \end{pmatrix}$$

(b) $g((x_1, x_2, x_3), (y_1, y_2, y_3)) = x_1 y_1 - x_1 y_2 + x_3 y_2 + x_2 y_2$

(c) $h((x_1, x_2, x_3), (y_1, y_2, y_3)) := \int_{-1}^1 (x_1 + x_2 x + x_3 x^2)(y_1 + y_2 x + y_3 x^2) dx$

2. Berechnen Sie die Gramsche Matrix von (v_1, v_2, v_3) bezüglich des Standardskalarproduktes auf \mathbb{R}^3 für

$$v_1 = (1, 1, 2), \quad v_2 = (0, 1, 3), \quad v_3 = (1, -2, 1).$$

3. Berechnen Sie den Winkel zwischen f und g , wobei

$$f = \cos x + \sin x + 3 \sin 2x$$

$$g = \cos 2x + \sin 2x + 3 \sin 3x$$

und das Skalarprodukt $\langle f, g \rangle = \int_0^{2\pi} f(x)g(x) dx$ zugrunde liegt.

4. Es seien $(V_1, g_1), (V_2, g_2)$ euklidische Vektorräume. Definieren Sie ein Skalarprodukt g auf $V_1 \times V_2$, so dass $V_1 \times 0$ und $0 \times V_2$ orthogonale Unterräume von $V_1 \times V_2$ sind.
5. Es sei V ein euklidischer Vektorraum mit Orthonormalbasis (v_1, \dots, v_4) .
Es sei U der von $v_1 + v_2$ und $v_1 + v_2 - v_3$ aufgespannte Untervektorraum.
Bestimmen Sie eine Basis von $U^\perp \subset V$.
6. Es sei U der von $(1, 1, 1, 0, 0, 0)$, $(0, 1, 1, 1, 0, 0)$ und $(1, 0, 1, 0, 1, 0)$ aufgespannte Untervektorraum von \mathbb{R}^6 .
- (a) Finden Sie eine Orthonormalbasis von U bzgl. des Standardskalarproduktes.
- (b) Berechnen Sie die orthogonale Projektion von $(1, 1, 1, 1, 1, 1)$ auf U .
7. Es sei (Ω, p) ein endlicher Wahrscheinlichkeitsraum, $\Omega = \{x_1, \dots, x_n\}$, $p : \Omega \rightarrow \mathbb{R}$, $p(x_i) > 0$, $\sum_{i=1}^n p(x_i) = 1$.
- Es sei $\mathcal{V} = \text{Abb}(\Omega, \mathbb{R})$ der Raum der Zufallsvariablen.

$$E : \mathcal{V} \longrightarrow \mathbb{R}$$

sei das Funktional $E(f) := \sum_{x \in \Omega} p(x)f(x)$.

(a) Zeigen Sie: Mit $\langle f, g \rangle := E(fg)$ ist \mathcal{V} ein euklidischer Vektorraum.

(b) Für eine Teilmenge $A \subset \Omega$ setzt man $p(A) := \sum_{x \in A} p(x)$.

Zwei Zufallsvariable $f, g \in \mathcal{V}$ heißen **unabhängig**, wenn für alle $a, b \in \mathbb{R}$ gilt

$$p(f^{-1}(a) \cap g^{-1}(b)) = p(f^{-1}(a))p(g^{-1}(b)).$$

Zeigen Sie: Sind $f, g \in \mathcal{V}$ mit $E(f) = E(g) = 0$ und sind f, g unabhängig, so sind f, g orthogonal.

(c) $\mathcal{V}_0 = \{f \in \mathcal{V} \mid E(f) = 0\}$ ist Untervektorraum von \mathcal{V} . Bestimmen Sie den orthogonalen Projektor von \mathcal{V} auf \mathcal{V}_0 .

18 Orthogonale Abbildungen

Es sei V ein n -dimensionaler euklidischer Vektorraum.

Definition 18.1 Für $v, w \in V$ heißt $d(v, w) := \|v - w\|$ der **euklidische Abstand** von v und w .

Definition 18.2 Eine Abbildung $F : V \rightarrow V$ heißt **Isometrie** (oder Kongruenzabbildung) wenn gilt:

$$d(F(v), F(w)) = d(v, w).$$

Beispiel 18.3 Ist $a \in V$, so ist die Translation $T_a : V \rightarrow V$, $T_a(v) = v + a$ eine Isometrie.

Lemma 18.4 Ist $F : V \rightarrow V$ eine Isometrie mit Fixpunkt $b \in V$, also $F(b) = b$, so ist $G := T_{-b} \circ F \circ T_b$ eine Isometrie mit $G(0) = 0$ und es gilt für alle $v, w \in V$:

$$\langle G(v), G(w) \rangle = \langle v, w \rangle.$$

Beweis:

1. G ist Isometrie als Komposition von Isometrien. $G(0) = F(b) - b = 0$.
2. Für $v \in V$ gilt

$$\|v\| = d(v, 0) = d(G(v), G(0)) = d(G(v), 0) = \|G(v)\|.$$

Für $v, w \in V$ gilt

$$\|v - w\|^2 = \|v\|^2 - 2\langle v, w \rangle + \|w\|^2$$

und

$$\|G(v) - G(w)\|^2 = \|G(v)\|^2 - 2\langle G(v), G(w) \rangle + \|G(w)\|^2.$$

Da nun $\|G(v)\| = \|v\|$, $\|G(w)\| = \|w\|$ und $\|v - w\| = \|G(v) - G(w)\|$, folgt durch Vergleich $\langle v, w \rangle = \langle G(v), G(w) \rangle$. \square

Definition 18.5 Eine Abbildung $F : V \rightarrow V$ heißt **orthogonal** $\iff \forall v, w \in V : \langle F(v), F(w) \rangle = \langle v, w \rangle$.

Satz 18.6 Jede orthogonale Abbildung $F : V \rightarrow V$ ist ein Vektorraumisomorphismus und eine Isometrie.

Beweis: (a) Wir zeigen: F ist isometrisch mit $F(0) = 0$. Aus $\langle F(0), F(0) \rangle = \langle 0, 0 \rangle = 0$, folgt $F(0) = 0$. Für $v, w \in V$ gilt

$$\begin{aligned} \|F(v) - F(w)\|^2 &= \langle F(v), F(v) \rangle - 2\langle F(v), F(w) \rangle + \langle F(w), F(w) \rangle \\ &= \langle v, v \rangle - 2\langle v, w \rangle + \langle w, w \rangle = \|v - w\|^2. \end{aligned}$$

Also ist F eine Isometrie.

(b) Wir zeigen, dass $F : V \rightarrow V$ Vektorraumisomorphismus ist.

Sei (e_1, \dots, e_n) eine Orthonormalbasis von V . Da $\langle F(e_i), F(e_j) \rangle = \langle e_i, e_j \rangle$, ist auch $(F(e_1), \dots, F(e_n))$ eine Orthonormalbasis von V . Durch

$$G \left(\sum_{i=1}^n x_i e_i \right) := \sum_{i=1}^n x_i F(e_i)$$

ist somit ein Vektorraumisomorphismus $G : V \rightarrow V$ definiert.

Offensichtlich gilt:

Ist $v = \sum_{i=1}^n x_i e_i$, so ist $x_i = \langle v, e_i \rangle$ und somit

$$F(v) = \sum_{i=1}^n \langle F(v), F(e_i) \rangle F(e_i) = \sum_{i=1}^n \langle v, e_i \rangle F(e_i) = \sum_{i=1}^n x_i F(e_i) = G(v)$$

Damit ist $F = G$. □

Lemma 18.7 Es sei $F : V \rightarrow V$ orthogonale Abbildung und $\mathcal{B} = (v_1, \dots, v_n)$ eine Basis von V , G ihre Gramsche Matrix. Weiter sei A die Matrix von F bezüglich der Basis \mathcal{B} .

Dann gilt

$$G = A^t G A.$$

Ist \mathcal{B} eine Orthonormalbasis (also $G = E$), so gilt

$$E = A^t A,$$

d.h. die Spalten von A bilden eine Orthonormalbasis von \mathbb{R}^n .

Weiter gilt $\det A = \pm 1$.

Beweis:

$$\begin{aligned} g_{ij} &= \langle v_i, v_j \rangle = \langle F(v_i), F(v_j) \rangle \\ &= \left\langle \sum_k a_{ki} v_k, \sum_\ell a_{\ell j} v_\ell \right\rangle \\ &= \sum_{k,\ell} a_{ki} g_{k\ell} a_{\ell j} \end{aligned}$$

$$\implies G = A^t G A$$

$$\implies \det G = (\det A)^2 \det G \implies \det A = \pm 1 \quad \square$$

Aus $A^t A = E$ folgt, dass $A^t = A^{-1}$. Es gilt also auch $AA^t = E$.

Definition 18.8 Eine Matrix $A \in M_n(\mathbb{R})$ heißt **orthogonal** $\iff A^t A = E$.

Mit $O(n)$ wird die Menge aller orthogonalen $n \times n$ -Matrizen bezeichnet.

Mit $O(V)$ bezeichnen wir die Menge aller orthogonalen Abbildungen $F : V \rightarrow V$.

Ist eine Orthonormalbasis von V fixiert und bezeichnet $A = \mathcal{M}(F)$ die Matrix von F bezüglich dieser Basis, so erhält man eine Abbildung

$$\mathcal{M} : O(V) \longrightarrow O(n)$$

mit folgenden Eigenschaften

- a) $\mathcal{M}(id_V) = E$
- b) $\mathcal{M}(F \circ G) = \mathcal{M}(F)\mathcal{M}(G)$.

Die Mengen $O(n)$ und $O(V)$ mit der Matrizen-Multiplikation bzw. Komposition von Abbildungen als „Verknüpfung“ sind Beispiele von **Gruppen**. Der Begriff der Gruppe, den wir jetzt einführen wollen, ist ein fundamentaler Begriff der Algebra.

Definition 18.9 Eine **Gruppe** ist eine Menge G zusammen mit einer Verknüpfung $\cdot : G \times G \longrightarrow G$, $(x, y) \longmapsto x \cdot y$, so dass folgende Axiome erfüllt sind:

(G1) Es gilt das Assoziativ-Gesetz:

$$\forall x, y, z \in G : x \cdot (y \cdot z) = (x \cdot y) \cdot z$$

(G2) Es gibt ein neutrales Element:

$$\exists e \in G \quad \forall x \in G : e \cdot x = x \cdot e = x$$

(Wie wir schon früher gesehen haben, ist e eindeutig bestimmt.)

(G3) Jedes Element besitzt ein Inverses:

$$\forall x \in G \exists y \in G : x \cdot y = y \cdot x = e$$

(y ist ebenfalls durch x eindeutig bestimmt.)

y heißt das **Inverse** von x und wird mit x^{-1} bezeichnet.

Es ist üblich, statt $x \cdot y$ kurz xy zu schreiben.

Man darf Klammern in mehrfachen Produkten weglassen:

$$xyz = (xy)z = x(yz).$$

Allerdings darf man im allgemeinen nicht die Reihenfolge der Faktoren vertauschen. Eine Gruppe heißt **abelsch**, wenn das Kommutativ-Gesetz gilt:

$$\forall x, y \in G : xy = yx.$$

Abelsche Gruppen werden oft **additiv** geschrieben. Statt $x \cdot y$ schreibt man also $x + y$.

Natürlich schreibt man dann auch $-x$ an Stelle von x^{-1} . Das neutrale Element e wird dann mit 0 bezeichnet.

Ist $(K, +, \cdot)$ ein Körper, so ist $(K, +)$ eine abelsche Gruppe. $(K, +)$ heißt die **additive Gruppe** des Körpers.

Auch $(K^\times = K \setminus \{0\}, \cdot)$ ist eine abelsche Gruppe. K^\times heißt die **multiplikative Gruppe** von K .

Ist G eine Gruppe, $n \in \mathbb{Z}, x \in G$, so kann man die Potenz $x^n \in G$ einführen:

Man setzt $x^0 := e$.

Für $n > 0$ setzt man $x^n := x \cdot x^{n-1}$.

Für $n < 0$ setzt man $x^n := (x^{-n})^{-1}$.

Man zeigt leicht die Potenz-Rechenregeln

$$\begin{aligned}x^{n+m} &= x^n x^m, \\x^{nm} &= (x^n)^m\end{aligned}$$

für alle $n, m \in \mathbb{Z}, x \in G$.

Weiter gilt

$$(xy)^n = x^n y^n,$$

falls $xy = yx$.

Definition 18.10 Es sei G eine Gruppe, $H \subset G$ eine Teilmenge. H heißt **Untergruppe** von G , wenn gilt:

- (a) $e \in H$
- (b) $\forall x, y \in G : x, y \in H \implies xy \in H$
- (c) $\forall x \in G : x \in H \implies x^{-1} \in H$.

Eine Untergruppe ist mit der von G induzierten Multiplikation natürlich ebenfalls eine Gruppe.

Definition 18.11 Es seien G_1, G_2 Gruppen. Eine Abbildung $\varphi : G_1 \longrightarrow G_2$ heißt **Gruppenhomomorphismus**, wenn gilt:

$$\forall x, y \in G_1 : \varphi(xy) = \varphi(x)\varphi(y).$$

Es folgt dann auch $\varphi(e) = e$, denn: $e \cdot \varphi(e) = \varphi(e) = \varphi(e \cdot e) = \varphi(e) \cdot \varphi(e)$, also $e = \varphi(e)$.

Weiter gilt $\varphi(x^{-1}) = \varphi(x)^{-1}$, denn:

$$e = \varphi(e) = \varphi(xx^{-1}) = \varphi(x)\varphi(x^{-1})$$

und somit

$$\varphi(x)^{-1} = \varphi(x)^{-1}e = \varphi(x)^{-1}\varphi(x)\varphi(x^{-1}) = e\varphi(x^{-1}) = \varphi(x^{-1}).$$

Ein **Gruppenisomorphismus** ist ein bijektiver Gruppenhomomorphismus.

Zwei Gruppen G_1, G_2 heißen **isomorph**, wenn es einen Isomorphismus $\varphi : G_1 \longrightarrow G_2$ gibt.

Lemma und Definition 18.12 Es sei M eine Menge

$$S(M) = \{f : M \longrightarrow M \mid f \text{ bijektiv}\}$$

ist mit der Komposition von Abbildungen als Verknüpfung eine Gruppe.

$S(M)$ heißt die (volle) **symmetrische Gruppe** von M .

Ist $M = \{1, \dots, n\}$, so heißt $S_n := S(M)$ die symmetrische Gruppe n -ten Grades.

Die Elemente von S_n sind *Permutationen*. \square

Lemma und Definition 18.13 Es sei R ein Ring mit Eins. Die Menge der Einheiten in R ist mit der Multiplikation eine Gruppe, die Einheitengruppe von R . Sie wird mit R^\times bezeichnet.

$$R^\times = \{x \in R \mid \exists y \in R : xy = yx = 1\}.$$

Definition 18.14 a) Es sei V ein n -dimensionaler K -Vektorraum.

$$GL(V) := \text{End}(V)^\times = \{F : V \rightarrow V \mid F \text{ Isomorphismus}\}$$

heißt die **allgemeine lineare Gruppe** von V .

b) Es sei R ein kommutativer Ring mit Eins. $n \in \mathbb{N}, n > 0$.

$$GL_n(R) = M_n(R)^\times = \{A \in M_n(R) \mid A \text{ ist invertierbar in } M_n(R)\}$$

heißt die **allgemeine lineare Gruppe n -ter Ordnung** über R .

Lemma 18.15 a) $SL(V) = \{F \in GL(V) \mid \det F = 1\}$

ist eine Untergruppe von $GL(V)$. Sie heißt die **spezielle lineare Gruppe** von V .

b) $SL_n(R) = \{A \in GL_n(R) \mid \det A = 1\}$

ist eine Untergruppe von $GL_n(R)$ und heißt die **spezielle lineare Gruppe n -ter Ordnung** über R .

Beweis: zu a):

1. $id_V \in SL(V)$, denn $\det(id_V) = 1$.
2. $F, G \in SL(V) \Rightarrow \det(F \circ G) = \det F \det G = 1 \cdot 1 = 1 \Rightarrow F \circ G \in SL(V)$.
3. $F \in SL(V) \Rightarrow 1 = \det(F \circ F^{-1}) = \det F \cdot \det(F^{-1}) = \det(F^{-1}) \Rightarrow F^{-1} \in SL(V)$.

zu b) analog. \square

Ist $\mathcal{B} = (v_1, \dots, v_n)$ eine Basis von V , so ist die Abbildung

$$\mathcal{M} : \text{End}(V) \longrightarrow M_n(K),$$

die F die Matrix $A = \mathcal{M}(F)$ von F bezüglich der Basis \mathcal{B} zuordnet, ein K -Algebra-Isomorphismus und induziert einen Gruppenisomorphismus

$$\mathcal{M} : GL(V) \longrightarrow GL_n(K).$$

$GL(V)$ ist isomorph zu $GL_n(K)$,

$SL(V)$ ist isomorph zu $SL_n(K)$.

Lemma 18.16 Es sei V ein euklidischer Vektorraum der Dimension n .

Dann gilt:

$O(V)$ ist Untergruppe von $GL(V)$,

$O(n)$ ist Untergruppe von $GL_n(\mathbb{R})$.

Die Abbildung

$$\mathcal{M} : O(V) \longrightarrow O(n),$$

die einer orthogonalen Abbildung $F : V \rightarrow V$ die Matrix $A = \mathcal{M}(F)$ bezüglich einer festen Orthonormalbasis \mathcal{B} zuordnet, ist ein Isomorphismus.

$O(V)$ heißt die **orthogonale Gruppe von V** ,

$O(n)$ heißt die **reelle orthogonale Gruppe n -ter Ordnung**.

Wieder hat man die „speziellen“ Untergruppen

$$SO(V) := O(V) \cap SL(V),$$

$$SO(n) := O(n) \cap SL_n(\mathbb{R}).$$

□

Beispiel 18.17 a) $O(1) = \{a \in \mathbb{R} \mid a^2 = 1\} = \{\pm 1\}$,

$$SO(1) = \{1\}.$$

b) Wir wollen zeigen, dass

$$SO(2) = \left\{ \begin{pmatrix} a & -b \\ b & a \end{pmatrix} \mid a, b \in \mathbb{R} \text{ und } a^2 + b^2 = 1 \right\}.$$

Beweis. Sei $A = \begin{pmatrix} a & c \\ b & d \end{pmatrix} \in SO(2)$. Dann gilt also $A^t A = E$, d.h.

$$\begin{pmatrix} a \\ b \end{pmatrix}, \begin{pmatrix} c \\ d \end{pmatrix}$$

bilden eine Orthonormalbasis von \mathbb{R}^2 , also gilt $a^2 + b^2 = 1$ und $\begin{pmatrix} a \\ b \end{pmatrix} = \pm \begin{pmatrix} -b \\ a \end{pmatrix}$

und da außerdem $\det A = 1$, folgt $\begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} -b \\ a \end{pmatrix}$. □

Es ist eine leichte Übung, zu zeigen, dass die Menge $S^1 = \{u \in \mathbb{C} \mid |u| = 1\}$ eine Untergruppe der multiplikativen Gruppe \mathbb{C}^\times ist und dass die Abbildung

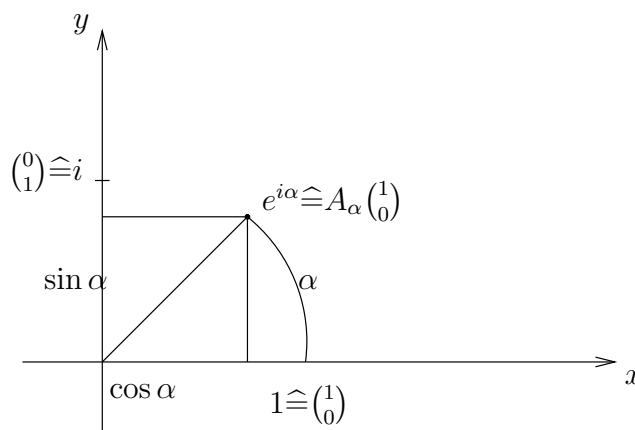
$$u = a + bi \longmapsto \begin{pmatrix} a & -b \\ b & a \end{pmatrix}$$

ein Gruppenisomorphismus $S^1 \rightarrow SO(2)$ ist. Die Abbildung $\mathbb{R} \rightarrow S^1$, $\alpha \mapsto e^{i\alpha} = \cos \alpha + i \sin \alpha$ ist ein surjektiver Gruppenhomomorphismus von der additiven Gruppe $(\mathbb{R}, +)$ auf S^1 , denn $e^{i(\alpha+\beta)} = e^{i\alpha} e^{i\beta}$.

Die Komposition $\mathbb{R} \rightarrow S^1 \rightarrow SO(2)$ ist der surjektive Gruppenhomomorphismus

$$\begin{aligned} \mathbb{R} &\longrightarrow SO(2) \\ \alpha &\longmapsto A_\alpha := \begin{pmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{pmatrix}. \end{aligned}$$

A_α ist die Drehung um den Winkel α im mathematisch positiven Sinn.



Die Gruppe $SO(2)$ ist abelsch und besteht aus den Drehungen um den Nullpunkt.

c) $O(2)$ ist nicht abelsch und besteht aus den Drehungen

$$A_\alpha = \begin{pmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{pmatrix}$$

und aus den Spiegelungen

$$S_\alpha = \begin{pmatrix} -\cos \alpha & \sin \alpha \\ \sin \alpha & \cos \alpha \end{pmatrix} = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} A_\alpha.$$

Man sieht leicht:

$$S_\alpha = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} A_{\frac{\alpha}{2}} A_{\frac{\alpha}{2}} = A_{-\frac{\alpha}{2}} \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} A_{\frac{\alpha}{2}} = (A_{\frac{\alpha}{2}})^{-1} S_0 A_{\frac{\alpha}{2}}.$$

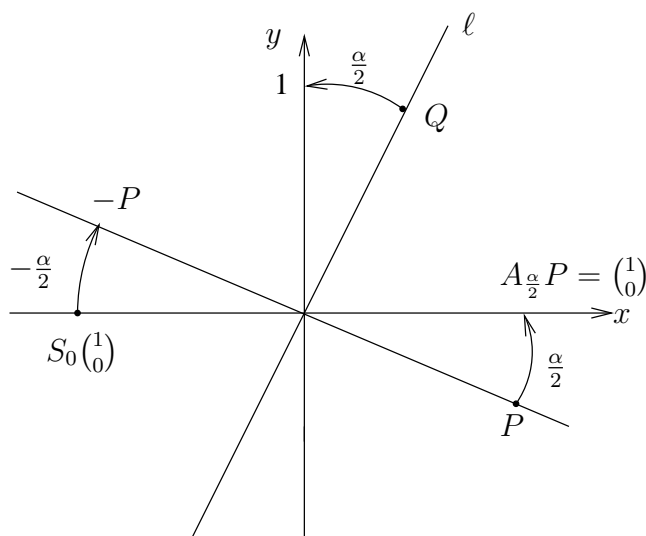
An dieser Darstellung kann man leicht die Eigenvektoren von S_α und damit die Spiegelachse ablesen:

Sei $P = (A_{\frac{\alpha}{2}})^{-1} \begin{pmatrix} 1 \\ 0 \end{pmatrix}$, $Q = (A_{\frac{\alpha}{2}})^{-1} \begin{pmatrix} 0 \\ 1 \end{pmatrix}$.

Dann ist

$$S_\alpha P = (A_{\frac{\alpha}{2}})^{-1} S_0 \begin{pmatrix} 1 \\ 0 \end{pmatrix} = -(A_{\frac{\alpha}{2}})^{-1} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = -P$$

und $S_\alpha Q = Q$.



Die Gerade durch 0 und Q ist die Spiegelachse ℓ von S_α .

Auch für $n \geq 3$ gilt

$$O(n) = SO(n) \cup S_0 \cdot SO(n),$$

wobei $S_0 = \begin{pmatrix} -1 & & & \\ & 1 & & \\ & & \ddots & \\ & & & 1 \end{pmatrix}$ die Spiegelung an der Koordinatenhyperebene $H = \{x \in \mathbb{R}^n \mid x_1 = 0\} \subset \mathbb{R}^n$ bezeichnet.

Eine Parametrisierung von $SO(n)$ ist für $n \geq 3$ nicht so einfach wie im Fall $n = 2$ zu bewerkstelligen. Die n^2 Komponenten a_{ij} einer $n \times n$ -Matrix unterliegen den 'Orthogonalitätsbedingungen'

$$\sum_{j=1}^n a_{ji}a_{jk} = \delta_{ik} \text{ für } 1 \leq i < k \leq n.$$

Das sind $\binom{n+1}{2}$ unabhängige Bedingungen und wir erwarten daher, dass $SO(n)$ von $\binom{n}{2} = n^2 - \binom{n+1}{2}$ reellen Parametern abhängt. Für $SO(3)$ werden wir dies näher ausführen.

Jetzt beweisen wir zunächst den Satz über die Normalform orthogonaler Abbildungen.

Wir benötigen den folgenden Hilfssatz.

Lemma 18.18 Es sei V ein n -dimensionaler \mathbb{R} -Vektorraum. Zu jedem Endomorphismus $F : V \rightarrow V$ gibt es einen F -invarianten Untervektorraum $U \subset V$ der Dimension 1 oder 2.

Beweis: (a) Hat F einen reellen Eigenwert, so hat V einen 1-dimensionalen F -invarianten Unterraum.

(b) F besitze keine reellen Eigenwerte.

mit $A_j \in SO(2)$ für $j = 1, \dots, k$ ($k \geq 0$).

Beweis: Induktion nach n .

Ist $n = 1$ oder $n = 2$, so ist nichts zu beweisen. Es sei $n \geq 3$ und die Zerlegungseigenschaft für orthogonale Abbildungen euklidischer Vektorräume der Dimension $\leq n-1$ schon bewiesen. Ist nun $\dim V = n$ und $F: V \rightarrow V$ orthogonal, so gibt es einen Untervektorraum U von V mit $1 \leq \dim U \leq 2$ und $F(U) = U$ (nach 18.18). Für U^\perp gilt dann $\dim U^\perp \leq n-1$. Da U orthogonal zu U^\perp ist und F orthogonal ist, ist $F(U)$ orthogonal zu $F(U^\perp)$: $u \in U, v \in U^\perp \Rightarrow \langle F(u), F(v) \rangle = \langle u, v \rangle = 0$.

Da $U = F(U)$, ist somit $F(U^\perp) \subset U^\perp$, U^\perp also F -invariant. Man erhält

$$F = F|_U \oplus F|_{U^\perp}.$$

Nach Induktionsvoraussetzung ist

$$F|_{U^\perp} = F_2 \oplus \dots \oplus F_s,$$

wobei $F_i: V_i \rightarrow V_i$ orthogonal ist und $\dim V_i \leq 2$. Mit $F_1 := F|_U$, $V_1 = U$ folgt dann

$$F = F_1 \oplus \dots \oplus F_s.$$

Die Matrixdarstellung folgt hieraus.

Man hat nur zu beachten, dass $F \in O(V) \setminus SO(V)$ im Fall $\dim V = 2$ bezüglich der Orthonormalbasis v_1, v_2 aus Eigenvektoren zum Eigenwert 1 bzw. -1 die Matrixdarstellung $\begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}$ besitzt. \square

Korollar 18.20 Es sei $n \geq 3$. Ist $A \in O(n)$, so gibt es eine Matrix $S \in SO(n)$, so dass

$$S^{-1}AS = \begin{pmatrix} 1 & & & & & & & 0 \\ & \ddots & & & & & & \\ & & 1 & & & & & \\ & & & -1 & & & & \\ & & & & \ddots & & & \\ & & & & & -1 & & \\ & & & & & & A_1 & \\ & & & & & & & \ddots \\ 0 & & & & & & & & A_k \end{pmatrix} \text{ mit } A_1, \dots, A_k \in SO(2)$$

Beweis: Nach 18.19 gibt es eine Matrix $T \in O(n)$, so dass

$$T^{-1}AT = \begin{pmatrix} 1 & & & & & & & & & & 0 \\ & \ddots & & & & & & & & & \\ & & 1 & & & & & & & & \\ & & & -1 & & & & & & & \\ & & & & \ddots & & & & & & \\ & & & & & -1 & & & & & \\ & & & & & & A_1 & & & & \\ & & & & & & & \ddots & & & \\ 0 & & & & & & & & & & A_k \end{pmatrix}$$

Ist $\det T = 1$, so ist $S = T$ wählbar; ist $\det T = -1$, so setze

$$S = T \begin{pmatrix} E_{n-2} & 0 \\ 0 & \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \end{pmatrix}.$$

Dann ist $S \in SO(n)$ und

$$S^{-1}AS = \begin{pmatrix} * & & 0 \\ 0 & \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} A_k \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \end{pmatrix}$$

falls $k \geq 1$, und $S^{-1}AS = T^{-1}AT$, falls $k = 0$. □

Korollar 18.21 (Euler)

Zu jeder Matrix $A \in SO(3)$ gibt es eine Matrix $S \in SO(3)$ und ein $\alpha \in \mathbb{R}$, so dass

$$S^{-1}AS = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos \alpha & -\sin \alpha \\ 0 & \sin \alpha & \cos \alpha \end{pmatrix}.$$

Jede Matrix $A \in SO(3)$ beschreibt also eine Drehung $A : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ um eine Gerade durch den Nullpunkt.

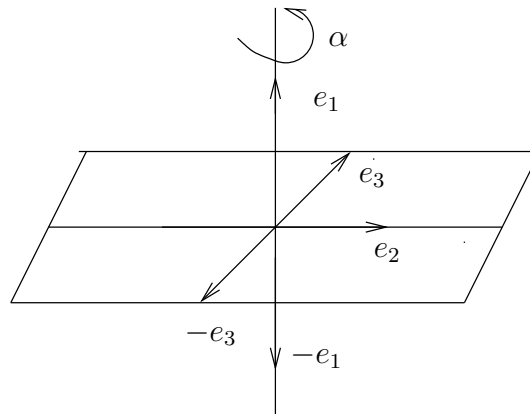
Beweis: Dies folgt aus 18.20 und $\det A = 1$. □

Beispiel 18.22 Eine Drehung F um die x_1 -Achse um den Winkel α wird bezüglich der Standardbasis (e_1, e_2, e_3) von \mathbb{R}^3 durch die Matrix $D = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos \alpha & -\sin \alpha \\ 0 & \sin \alpha & \cos \alpha \end{pmatrix}$ beschrieben.

Die Basis $(-e_1, e_2, -e_3)$ ist genauso orientiert wie (e_1, e_2, e_3) . Die Transformationsmatrix $S = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$ liegt in $SO(3)$ und es gilt für die Matrix \tilde{D} von F bzgl.

$(-e_1, e_2, -e_3)$

$$\tilde{D} = S^{-1}DS = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos(-\alpha) & -\sin(-\alpha) \\ 0 & \sin(-\alpha) & \cos(-\alpha) \end{pmatrix}$$



denn:

$$\begin{aligned} F(-e_1) &= -e_1 \\ F(e_2) &= (\cos \alpha)e_2 + (\sin \alpha)e_3 \\ &= (\cos(-\alpha))e_2 - (\sin \alpha)(-e_3) \\ &= (\cos(-\alpha))e_2 + (\sin(-\alpha))(-e_3) \\ F(-e_3) &= -F(e_3) = (\sin \alpha)e_2 - (\cos \alpha)e_3 \\ &= -(\sin(-\alpha))e_2 + (\cos(-\alpha))(-e_3). \end{aligned}$$

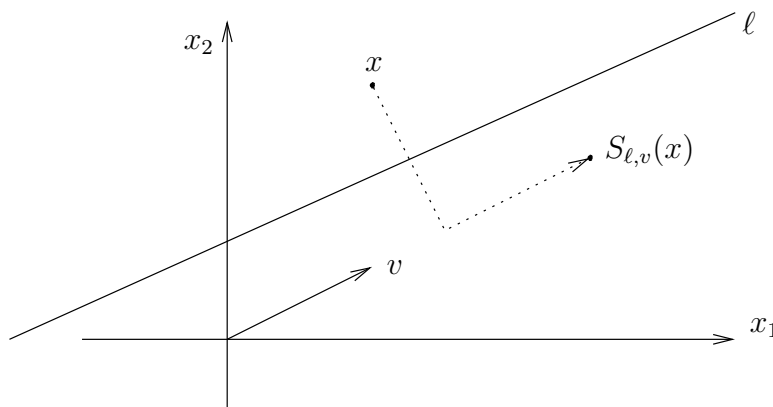
Übungen

- Es sei $\varphi : G \rightarrow H$ ein Gruppenhomomorphismus. Zeigen Sie:
 - $\{x \in G \mid \varphi(x) = e\}$ ist eine Untergruppe von G .
 - $\{\varphi(x) \mid x \in G\}$ ist eine Untergruppe von H .
- Ein zweidimensionaler euklidischer Vektorraum heißt euklidische Ebene. Zeigen Sie: Die einzigen Isometrien der euklidischen Ebene \mathbb{R}^2 sind die Translationen, Drehungen, Spiegelungen und Gleitspiegelungen.
 - Translationen sind die Abbildungen T_a mit $T_a(x) = x + a$.
 - Drehungen sind die Abbildungen $D_{a,\alpha}$ mit $D_{a,\alpha}(x) = A_\alpha(x - a) + a$, $A_\alpha = \begin{pmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{pmatrix}$.
 - Spiegelungen sind die Abbildungen S_ℓ mit $S_\ell(x) = S_{\ell_0}(x - a) + a$, wobei $\ell_0 \subset \mathbb{R}^2$ 1-dim. Untervektorraum, $\ell = \ell_0 + a$ Gerade, und $S_{\ell_0}(x_1 + x_2) = x_1 - x_2$ für $x_1 \in \ell_0$, $x_2 \in \ell_0^\perp$.

- Gleitspiegelungen $S_{\ell,v}$ sind die Abbildungen

$$S_{\ell,v} = T_v \circ S_\ell, \text{ wobei } v \in \mathbb{R}^2 \setminus \{0\}$$

und ℓ eine Gerade mit Richtungsvektor v ist.



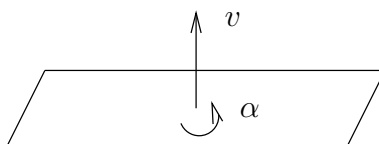
3. $H = \left\{ \begin{pmatrix} 1 & x \\ 0 & 1 \end{pmatrix} \mid x \in \mathbb{R} \right\}$ ist eine Untergruppe von $SL_2(\mathbb{R})$.

4. Es sei $S^2 = \{v \in \mathbb{R}^3 \mid \|v\| = 1\}$.

Untersuchen Sie die Abbildung

$$\Psi : S^2 \times (0, 2\pi) \longrightarrow SO(3),$$

die einem Paar (v, α) die Drehung um die Achse $\mathbb{R}v$ um den Winkel α im Gegenuhrzeigersinn zuordnet.



Zeigen Sie, dass $\Psi : S^2 \times (0, 2\pi) \longrightarrow SO(3) \setminus \{E\}$ surjektiv ist und dass jede Matrix $A \in SO(3)$, $A \neq E$, genau zwei Urbilder unter der Abbildung Ψ hat.

5. Es sei $A \in SO(3) \setminus \{E\}$. Welche geometrische Bedeutung haben

$$\ker(A - E) \text{ und } \frac{1}{2}(\text{Spur}(A) - 1)?$$

6. Es sei V ein n -dimensionaler euklidischer Vektorraum. Für $a \in V \setminus \{0\}$, sei $\sigma_a : V \longrightarrow V$ definiert durch

$$\sigma_a(v) = v - 2 \frac{\langle v, a \rangle}{\langle a, a \rangle} a$$

- (a) Zeigen Sie: σ_a ist orthogonale Abbildung.
 (b) Finden Sie die in 18.19 beschriebene Matrix für die Abbildung σ_a .

- (c) Es sei $S : V \rightarrow V$ eine orthogonale Abbildung. Zeigen Sie, dass $S^{-1} \circ \sigma_a \circ S = \sigma_b$ für ein $b \in V \setminus \{0\}$, welches?
- (d) Seien $a_1, a_2 \in V \setminus \{0\}$ mit $\|a_1\| = \|a_2\|$. Setze $a := a_1 - a_2$. Berechnen Sie $\sigma_a(a_1)$.
7. Es sei V ein n -dim. euklidischer Vektorraum. Zeigen Sie: Jede orthogonale Abbildung $F : V \rightarrow V$ ist die Komposition von m ($m \leq n$) Spiegelungen $\sigma_{a_1}, \dots, \sigma_{a_m}$.
- Hinweis: Ist $F(a) \neq a$, so findet man eine Spiegelung σ , so dass $(\sigma \circ F)(a) = a$. Iteriere!

19 Unitäre Vektorräume

Die Geometrie der euklidischen Ebene \mathbb{R}^2 lässt sich algebraisch elegant behandeln, wenn man \mathbb{R}^2 als komplexe Zahlenebene \mathbb{C} auffasst. Ein Paar (x, y) definiert die komplexe Zahl $z = x + iy$.

Kann man das Skalarprodukt auf \mathbb{R}^2 komplex interpretieren? Das ist einfach:

Sei $z = x + iy$, $w = u + iv$, $(x, y), (u, v) \in \mathbb{R}^2$.

Dann gilt

$$(x, y) \cdot (u, v) = xu + yv$$

und

$$zw = (xu - yv) + i(xv + yu).$$

Der Realteil von zw ist „fast“ dasselbe wie das Skalarprodukt $(x, y) \cdot (u, v)$. Das Vorzeichen vor yv ist falsch. Durch folgenden Trick wird dies repariert: Betrachte einfach

$$z\bar{w} = (xu + yv) + i(yu - xv).$$

Jetzt gilt in der Tat

$$\operatorname{Re}(z\bar{w}) = (x, y) \cdot (u, v).$$

Wir haben also das reelle Skalarprodukt auf die Multiplikation komplexer Zahlen zurückgeführt. $z\bar{w}$ enthält aber noch mehr Information, den Imaginärteil:

$$-\operatorname{Im}(z\bar{w}) = xv - yu = \det \begin{pmatrix} x & y \\ u & v \end{pmatrix}.$$

Diese Zahl gibt den (orientierten) Flächeninhalt des von (x, y) und (u, v) aufgespannten Parallelogramms in $\mathbb{R}^2 = \mathbb{C}$ an.

Allgemeiner kann man \mathbb{R}^{2n} als \mathbb{C}^n auffassen und für $z, w \in \mathbb{C}^n$, $z = x + iy$, $w = u + iv$ mit $u, v, x, y \in \mathbb{R}^n$ das so genannte unitäre Skalarprodukt

$$\langle z, w \rangle := z^t \bar{w} = \sum_{j=1}^n z_j \bar{w}_j$$

einführen. Es gilt

$$\operatorname{Re}\langle z, w \rangle = \sum_{j=1}^n x_j u_j + y_j v_j = (x, y) \cdot (u, v)$$

ist das Standard-Skalarprodukt auf \mathbb{R}^{2n} .

Ab jetzt sei in diesem Abschnitt V ein \mathbb{C} -Vektorraum.

Definition 19.1 Eine Abbildung

$$h : V \times V \longrightarrow \mathbb{C}$$

heißt **Sesquilinearform** auf V : \iff

(a) h ist \mathbb{C} -linear in der ersten Variablen, d.h.

$$h(av + bw, u) = ah(v, u) + bh(w, u)$$

für alle $a, b \in \mathbb{C}$, $v, w, u \in V$.

(b) h ist \mathbb{C} -antilinear in der zweiten Variablen, d.h.

$$h(u, av + bw) = \bar{a}h(u, v) + \bar{b}h(u, w)$$

für alle $a, b \in \mathbb{C}$, $u, v, w \in V$.

Ist $\dim_{\mathbb{C}} V = n$ und $\mathcal{B} = (v_1, \dots, v_n)$ eine Basis von V , so heißt $H = (h_{jk})$ mit $h_{jk} = h(v_j, v_k)$ die Matrix von h bezüglich \mathcal{B} oder auch die Gramsche Matrix von (v_1, \dots, v_n) bezüglich h .

h ist durch H vollständig bestimmt, denn für $v = \sum_{j=1}^n z_j v_j$, $w = \sum_{j=1}^n w_j v_j$, $z_j, w_j \in \mathbb{C}$ gilt

$$h(v, w) = \sum_{j,k=1}^n z_j \bar{w}_k h_{jk} = z^t H \bar{w},$$

wobei $z = [z_1, \dots, z_n]$, $w = [w_1, \dots, w_n] \in \mathbb{C}^n$ und $\bar{w} := [\bar{w}_1, \dots, \bar{w}_n]$.

Wie im Reellen gilt

Lemma 19.2 (Transformationsformel)

Es sei $h : V \times V \rightarrow \mathbb{C}$ eine Sesquilinearform auf V , H, H' seien die Matrizen von h bezüglich der Basen $\mathcal{B} = (v_1, \dots, v_n)$, $\mathcal{B}' = (v'_1, \dots, v'_n)$ und es sei $S = (s_{ij})$ mit $v'_j = \sum_{i=1}^n s_{ij} v_i$.

Dann gilt

$$H' = S^t H \bar{S},$$

wobei $\bar{S} = (\bar{s}_{ij})$ die konjugiert komplexe Matrix zu S ist. □

Definition 19.3 a) Eine **hermitesche Form** auf V ist eine Sesquilinearform $h : V \times V \rightarrow \mathbb{C}$ mit der Eigenschaft

$$h(v, w) = \overline{h(w, v)} \text{ für alle } v, w \in V.$$

Die Matrix $H = (h_{ij})$ von h bezüglich einer Basis von V ist dann eine **hermitesche Matrix**, d.h. es gilt

$$h_{ij} = \overline{h_{ji}} \text{ kurz: } H = \bar{H}^t.$$

Für eine hermitesche Form ist $h(v, v)$ reell, denn $h(v, v) = \overline{h(v, v)}$.

b) Eine hermitesche Form $h : V \times V \rightarrow \mathbb{C}$ heißt **positiv definit**: $\iff \forall v \in V \setminus \{0\} : h(v, v) > 0$.

c) Eine hermitesche Matrix H heißt **positiv definit**, wenn für alle $z \in \mathbb{C}^n$, $z \neq 0$ gilt:

$$z^t H \bar{z} > 0.$$

- d) Ein **unitäres Skalarprodukt** auf V ist eine positiv definite hermitesche Form auf V .
- e) Ein endlich dimensionaler \mathbb{C} -Vektorraum mit einem unitären Skalarprodukt heißt **unitärer Vektorraum**.

Für das unitäre Skalarprodukt schreibt man gewöhnlich wieder $\langle u, v \rangle$ statt $h(u, v)$.

Beispiel 19.4 a) $(\mathbb{C}^n, \langle \cdot, \cdot \rangle)$, $\langle z, w \rangle = z^t \bar{w}$,
das unitäre Standard-Skalarprodukt auf \mathbb{C}^n .

- b) Es sei V der von den komplexen Funktionen

$$f_n : [0, 2\pi] \longrightarrow \mathbb{C}, \quad f_n(x) = e^{inx} = \cos(nx) + i \sin(nx), \quad n \in \mathbb{Z}$$

aufgespannte \mathbb{C} -Vektorraum.

(Hier ist natürlich $\dim_{\mathbb{C}} V = \infty$.)

Für $f, g \in V$ setzt man

$$\langle f, g \rangle := \int_0^{2\pi} f(x) \overline{g(x)} dx.$$

Dies ist eine positiv definite hermitesche Form auf V .

Die Elemente von V heißen komplexe Fourierpolynome.

Definition 19.5 Es sei V ein unitärer Vektorraum.

Für $v \in V$ setzt man

$$\|v\| = \sqrt{\langle v, v \rangle}.$$

$\|v\|$ heißt die (unitäre) Norm von v .

Für $\lambda \in \mathbb{C}$, $v \in V$ gilt

$$\|\lambda v\| = \sqrt{\langle \lambda v, \lambda v \rangle} = \sqrt{\lambda \bar{\lambda} \langle v, v \rangle} = |\lambda| \|v\|.$$

Satz 19.6 (Cauchy-Schwarzsche Ungleichung)

Es sei V ein unitärer Vektorraum. Dann gilt

$$|\langle v, w \rangle| \leq \|v\| \|w\| \quad \text{für alle } v, w \in V.$$

Gleichheit gilt genau dann, wenn v, w \mathbb{C} -linear abhängig sind.

Beweis: Zum Beweis wollen wir den reellen Fall verwenden. V kann auch als \mathbb{R} -Vektorraum aufgefasst werden, weil $\mathbb{R} \subset \mathbb{C}$. Dann ist $\operatorname{Re}\langle \cdot, \cdot \rangle : V \times V \longrightarrow \mathbb{R}$ ein euklidisches Skalarprodukt auf V mit derselben Norm wie $\langle \cdot, \cdot \rangle$. Nach der reellen Cauchy-Schwarzschen Ungleichung gilt

$$\operatorname{Re}\langle v, w \rangle \leq \|v\| \|w\| \quad \text{für alle } v, w \in V.$$

Es gilt $\langle v, w \rangle = u |\langle v, w \rangle|$ mit $u \in S^1$, also $u\bar{u} = 1$ und somit

$$|\langle v, w \rangle| = u\bar{u} |\langle v, w \rangle| = \bar{u} \langle v, w \rangle = \langle v, uw \rangle.$$

Insbesondere ist $\langle v, uw \rangle$ reell und nicht negativ. Es folgt

$$\begin{aligned} |\langle v, w \rangle| &= \langle v, uw \rangle = |\operatorname{Re}\langle v, uw \rangle| \leq \|v\| \|uw\| \\ &= \|v\| \|u\| \|w\| = \|v\| \|w\|. \end{aligned}$$

Sind v, w \mathbb{C} -linear abhängig, so gilt offensichtlich Gleichheit. Gilt $|\langle v, w \rangle| = \|v\| \|w\|$, so gilt auch $|\operatorname{Re}\langle v, uw \rangle| = \|v\| \|uw\|$, also sind v und uw \mathbb{R} -linear abhängig, und somit sind v und w \mathbb{C} -linear abhängig. \square

Es sei V ein n -dimensionaler unitärer Vektorraum. Ohne Mühe kann man Definition 17.12, Satz 17.13, Definition 17.15, Lemma 17.16, Satz 17.17, Definition 17.18, Lemma 17.19 und Satz 17.20 übertragen.

Auch das Analogon zu Satz 17.22 ist richtig: Eine hermitesche Matrix H ist genau dann positiv definit, wenn die Hauptminoren $\det H_k$, $k = 1, \dots, n$ positiv sind.

Beispiel 19.7 a) $H = \begin{pmatrix} 1 & -i \\ i & 1 \end{pmatrix}$ ist hermitesch, denn

$$\bar{H} = \begin{pmatrix} 1 & i \\ -i & 1 \end{pmatrix}, \quad \bar{H}^t = H.$$

Aber H ist nicht positiv definit, weil $\det H = 0$.

b) $H = \begin{pmatrix} 1 & 1-i \\ 1+i & 3 \end{pmatrix}$ ist positiv definit, denn:

$$H_1 = 1 > 0, \quad \det H = 3 - (1-i)(1+i) = 1 > 0.$$

Beispiel 19.8 Sei V das Beispiel 19.4 (b). Wir betrachten den $(2N+1)$ -dimensionalen \mathbb{C} -Untervektorraum V_N , der von

$$f_n(x) = e^{inx}, \quad |n| \leq N$$

erzeugt wird.

Es sei $\varphi_n(x) = \frac{1}{\sqrt{2\pi}} f_n(x)$.

Behauptung: $\{\varphi_n \mid |n| \leq N\}$ ist eine Orthonormalbasis von V_N .

Beweis:

$$\|\varphi_n\|^2 = \int_0^{2\pi} \varphi_n(x) \overline{\varphi_n(x)} dx = \frac{1}{2\pi} \int_0^{2\pi} dx = 1.$$

Für $n \neq m$ ist

$$\begin{aligned} \langle \varphi_n, \varphi_m \rangle &= \frac{1}{2\pi} \int_0^{2\pi} e^{inx} \overline{(e^{imx})} dx = \\ &= \frac{1}{2\pi} \int_0^{2\pi} e^{inx} e^{-imx} dx = \frac{1}{2\pi} \int_0^{2\pi} e^{i(n-m)x} dx = \\ &= \frac{1}{2\pi} \left[\frac{1}{i(n-m)} e^{i(n-m)x} \right]_0^{2\pi} = 0. \quad \square \end{aligned}$$

Sei $f \in V$. f besitzt die orthogonale Projektion

$$f_N := \sum_{n=-N}^N \langle f, \varphi_n \rangle \varphi_n \in V_N$$

auf den Unterraum V_N .

$$\langle f, \varphi_n \rangle = \frac{1}{\sqrt{2\pi}} \int_0^{2\pi} f(x) e^{-inx} dx$$

heißt der n -te **komplexe Fourierkoeffizient** von f .

Da f_N orthogonal zu $f - f_N$, gilt

$$\langle f, f_N \rangle = \langle f - f_N, f_N \rangle + \langle f_N, f_N \rangle = \|f_N\|^2 = \sum_{n=-N}^N \langle f, \varphi_n \rangle^2$$

und somit

$$\begin{aligned} \|f - f_N\|^2 &= \langle f - f_N, f - f_N \rangle = \langle f - f_N, f \rangle \\ &= \|f\|^2 - \langle f_N, f \rangle = \|f\|^2 - \sum_{n=-N}^N \langle f, \varphi_n \rangle^2, \end{aligned}$$

also gilt $\|f\|^2 \geq \sum_{n=-N}^N \langle f, \varphi_n \rangle^2$ für alle N .

Die Fortführung dieser Überlegungen gehört in die Analysis.

Übungen

1. Es sei U der von $(i, 0, 1)$ und $(1, 1, i)$ aufgespannte komplexe Unterraum von \mathbb{C}^3 . Bestimmen Sie U^\perp .
2. Sind $v = (3i+1, 1-i, i+1)$ und $w = (-1+7i, 3-i, 1+3i)$ \mathbb{C} -linear abhängig. Berechnen Sie $\langle v, w \rangle, \|v\|, \|w\|$.
3. In \mathbb{C}^4 sei das unitäre Standard-Skalarprodukt gegeben. Wenden Sie das Schmidtsche Orthonormalisierungsverfahren auf (v_1, v_2) an, wobei

$$\begin{aligned} v_1 &= (i, 1, 0, i+1) \\ v_2 &= (i, i, i+1, 1). \end{aligned}$$

Interpretieren Sie das Ergebnis.

4. Berechnen Sie eine Orthonormalbasis von \mathbb{C}^2 bezüglich des unitären Skalarprodukts h mit der Matrix $H = \begin{pmatrix} 1 & i+1 \\ 1-i & 3 \end{pmatrix}$ bzgl. der Standardbasis von \mathbb{C}^2 .

20 Die unitäre Gruppe

Definition 20.1 (a) Es sei V ein n -dimensionaler unitärer Vektorraum. $F \in \text{End}V$ heißt **unitär** $\iff \forall v, w \in V : \langle F(v), F(w) \rangle = \langle v, w \rangle$.

Die unitären Endomorphismen sind offensichtlich Isomorphismen (denn wegen $\|F(v)\| = \|v\|$ sind unitäre Endomorphismen injektiv, also bijektiv) und bilden eine Untergruppe

$$U(V) \subset GL(V).$$

$U(V)$ heißt die **unitäre Gruppe** von V .

(b) Eine Matrix $A \in M_n(\mathbb{C})$ heißt **unitär** $\iff A\bar{A}^t = E$.

Mit $U(n)$ sind die Menge aller unitären $n \times n$ -Matrizen bezeichnet.

Offensichtlich ist $U(n) \subset GL_n(\mathbb{C})$, denn ist $A\bar{A}^t = E$, so ist A invertierbar mit $A^{-1} = \bar{A}^t$. Außerdem folgt aus dem Determinantenmultiplikationssatz, dass $\det A$ eine unimodulare komplexe Zahl ist: $\det A \det \bar{A} = 1$.

$U(n)$ ist eine Untergruppe von $GL_n(\mathbb{C})$ denn:

(a) $E \in U(n)$

(b) $A, B \in U(n) \Rightarrow AB(\overline{AB})^t = AB(\overline{A}\overline{B})^t = AB\bar{B}^t\bar{A}^t = A\bar{A}^t = E$, also ist $AB \in U(n)$.

(c) Ist $A \in U(n)$ so ist $A^{-1} = \bar{A}^t$ und somit ist $A^{-1}(\overline{A^{-1}})^t = A^{-1}(\overline{\bar{A}^t})^t = A^{-1}A = E$, also $A^{-1} \in U(n)$.

$U(n)$ heißt die **unitäre Gruppe** n -ter Ordnung.

$SU(n) = \{A \in U(n) \mid \det A = 1\}$ ist eine Untergruppe von $U(n)$. Sie heißt die **spezielle unitäre Gruppe** n -ter Ordnung.

Es sei $\mathcal{B} = (v_1, \dots, v_n)$ eine Orthonormalbasis von V . Dann ist

$$F \mapsto \mathcal{M}(F) = A = (\text{Matrix von } F \text{ bezüglich } \mathcal{B})$$

ein Gruppenisomorphismus.

$$\mathcal{M} : U(V) \rightarrow U(n).$$

Es sei $A = (a_{ij})$ mit $F(v_j) = \sum_{k=1}^n a_{kj}v_k$.

Dann gilt

$$\begin{aligned} \delta_{ij} = \langle v_i, v_j \rangle &= \langle F(v_i), F(v_j) \rangle = \sum_{l,k=1}^n a_{li}\bar{a}_{kj} \langle v_l, v_k \rangle \\ &= \sum_{l,k=1}^n a_{li}\bar{a}_{kj} \delta_{lk} = \sum_{k=1}^n a_{ki}\bar{a}_{kj}, \text{ d.h.} \end{aligned}$$

$A^t\bar{A} = E$, also auch $\bar{A}^t A = E$, d.h. $A \in U(n)$.

Beispiel 20.2

(a) $U(1) = \{a \in \mathbb{C} \mid a\bar{a} = 1\} = S^1 \cong SO(2)$.

(b) Die Gruppe $SU(2)$ ist wichtig in der Physik. Sie heißt auch **Spingruppe**.

Wir zeigen, dass

$$SU(2) = \left\{ \begin{pmatrix} a & -\bar{b} \\ b & \bar{a} \end{pmatrix} \mid a, b \in \mathbb{C} \text{ und } |a|^2 + |b|^2 = 1 \right\}$$

Beweis: Es sei $A = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} \in SU(2)$.

Es gilt dann $A^t \bar{A} = E$ und das heißt

$$\begin{pmatrix} \alpha & \gamma \\ \beta & \delta \end{pmatrix} \begin{pmatrix} \bar{\alpha} & \bar{\beta} \\ \bar{\gamma} & \bar{\delta} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Das ergibt die drei Bedingungen

- (i) $|\alpha|^2 + |\gamma|^2 = 1$
- (ii) $\alpha\bar{\beta} + \gamma\bar{\delta} = 0$
- (iii) $|\beta|^2 + |\delta|^2 = 1$

Weiter gilt $\det A = 1$, d.h.

- (iv) $\alpha\delta - \beta\gamma = 1$.

Multipliziert man (i) mit $\bar{\beta}$, so erhält man

$$\bar{\beta}\alpha\bar{\alpha} + \bar{\beta}\gamma\bar{\gamma} = \bar{\beta}.$$

Ersetzt man hier $\bar{\beta}\alpha$ durch $-\gamma\bar{\delta}$ (was nach (iii) erlaubt ist), so ergibt sich

$$-\gamma\bar{\delta}\bar{\alpha} + \bar{\beta}\gamma\bar{\gamma} = \bar{\beta},$$

also

$$-\gamma(\overline{\alpha\delta - \beta\gamma}) = \bar{\beta}$$

und somit nach (iv):

$$\bar{\beta} = -\gamma, \text{ also } \beta = -\bar{\gamma}.$$

Setzt man dies in (ii) ein, so erhält man

$$-\alpha\gamma + \gamma\bar{\delta} = 0.$$

Ist $\gamma \neq 0$, so folgt $\alpha = \bar{\delta}$, also $\delta = \bar{\alpha}$ und somit hat

$$A = \begin{pmatrix} \alpha & -\bar{\gamma} \\ \gamma & \bar{\alpha} \end{pmatrix}$$

die gewünschte Gestalt.

Ist $\gamma = 0$, so lauten die Bedingungen (i) - (iv):

$$|\alpha| = 1, \quad \beta = 0, \quad |\delta| = 1, \quad \alpha\delta = 1.$$

Es folgt $\delta = \alpha^{-1} = \bar{\alpha}$ und somit hat auch in diesem Fall

$$A = \begin{pmatrix} \alpha & 0 \\ 0 & \bar{\alpha} \end{pmatrix}$$

die gewünschte Form.

Aus dieser Beschreibung von $SU(2)$ ist sofort ersichtlich, dass $SU(2)$ die Gestalt der dreidimensionalen Einheitskugel

$$S^3 = \{x = (x_1, x_2, x_3, x_4) \in \mathbb{R}^4 \mid x_1^2 + x_2^2 + x_3^2 + x_4^2 = 1\}$$

hat. Die Abbildung

$$\Phi : S^3 \rightarrow SU(2) \quad \text{mit} \quad \Phi(x) = \begin{pmatrix} x_1 + ix_2 & -x_3 + ix_4 \\ x_3 + ix_4 & x_1 - ix_2 \end{pmatrix} \quad \text{ist bijektiv.}$$

Im nächsten Abschnitt werden wir den Zusammenhang zwischen den Gruppen $SU(2)$ und $SO(3)$ genauer untersuchen.

Die Klassifikation der unitären Endomorphismen eines unitären Raumes ist einfacher als die der orthogonalen Endomorphismen eines euklidischen Vektorraums. Es gilt

Satz 20.3 Es sei (V, \langle, \rangle) ein n -dimensionaler unitärer Raum und $F : V \rightarrow V$ ein unitärer Endomorphismus. Dann gilt

- (a) Die Eigenwerte $\lambda \in \mathbb{C}$ von F haben den Absolutbetrag 1, sind also unimodulare komplexe Zahlen.
- (b) Es gibt eine Orthonormalbasis $\mathcal{B} = (v_1, \dots, v_n)$ aus Eigenvektoren von F .

Beweis: zu (a). Ist $F(v) = \lambda(v)$, so gilt $\langle v, v \rangle = \langle F(v), F(v) \rangle = \langle \lambda v, \lambda v \rangle = \lambda \bar{\lambda} \langle v, v \rangle$. Da $v \neq 0$ ist, gilt $\lambda \bar{\lambda} = 1$, d.h. $|\lambda| = 1$.

zu (b). Induktion nach n . Für $n = 1$ ist nichts zu beweisen.

Induktionsschluss $n - 1 \rightarrow n (n \leq 2)$: Es sei $v_1 \in V$ ein normierter Eigenvektor von F . $W = \mathbb{C}v_1$ ist ein eindimensionaler F -invarianter Unterraum. W^\perp ist $(n - 1)$ -dimensionaler Untervektorraum von V mit $V = W \oplus W^\perp$. W^\perp ist ebenfalls F -invariant, denn für $w \in W^\perp$ ist $0 = \langle w, v_1 \rangle = \langle F(w), F(v_1) \rangle = \langle F(w), \lambda_1 v_1 \rangle = \bar{\lambda}_1 \langle F(w), v_1 \rangle$, wobei λ_1 Eigenwert und $F(v_1) = \lambda_1 v_1$. Da $\lambda_1 \neq 0$, folgt somit

$$\langle F(w), v_1 \rangle = 0, \quad \text{d.h. } F(w) \in W^\perp.$$

Nach Induktionsvoraussetzung gibt es eine Orthonormalbasis (v_2, \dots, v_n) von $F|_{W^\perp} : W^\perp \rightarrow W^\perp$ aus Eigenvektoren von $F|_{W^\perp}$. Das sind natürlich auch Eigenvektoren von F und $\mathcal{B} = (v_1, \dots, v_n)$ ist eine Orthonormalbasis aus Eigenvektoren von F .

□

Korollar 20.4 Zu jeder unitären Matrix $A \in U(n)$ gibt es eine unitäre Matrix $B \in U(n)$, so dass

$$BAB^t = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix} \text{ mit}$$

unimodularen komplexen Zahlen $\lambda_1, \dots, \lambda_n \in S^1$.

Bemerkung 20.5

Um eine unitäre Matrix $A \in U(n)$ auf ihre Normalform zu bringen, geht man folgendermaßen vor: Zunächst bestimmt man die Nullstellen des charakteristischen Polynoms P_A .

Seien $\lambda_1, \dots, \lambda_m$ die verschiedenen Nullstellen. Jetzt berechne man eine Orthonormalbasis $(v_1^{(j)}, \dots, v_{k_j}^{(j)})$ des Eigenraums $E(\lambda_j)$ von λ_j . Dann ist die Matrix

$$S = (v_1^{(1)}, \dots, v_{k_1}^{(1)}, \dots, v_1^{(m)}, \dots, v_{k_m}^{(m)})$$

eine unitäre Matrix und nach Konstruktion der $v_l^{(j)}$ gilt die Matrixgleichung

$$AS = S \begin{pmatrix} \lambda_1 E_{k_1} & & 0 \\ & \ddots & \\ 0 & & \lambda_m E_{k_m} \end{pmatrix},$$

also

$$S^{-1}AS = \left(\begin{array}{cccc} \lambda_1 & & & 0 \\ & \ddots & & \\ & & \lambda_1 & \\ & & & \ddots \\ 0 & & & & \lambda_1 & & \\ & & & & & \ddots & \\ & & & & & & \lambda_1 \end{array} \right) \left. \begin{array}{l} \left. \vphantom{\begin{matrix} \lambda_1 \\ \ddots \\ \lambda_1 \\ \ddots \\ \lambda_1 \end{matrix}} \right\} k_1\text{-mal} \\ \left. \vphantom{\begin{matrix} \lambda_1 \\ \ddots \\ \lambda_1 \\ \ddots \\ \lambda_1 \end{matrix}} \right\} k_m\text{-mal} \end{array} \right.$$

Beispiel 20.6 Die Normalform einer Matrix $A \in SU(2)$ ist

$$\begin{pmatrix} \lambda & 0 \\ 0 & \bar{\lambda} \end{pmatrix}$$

wobei $\lambda \in S^1$. Es ist nämlich $\det A = 1$ und folglich sind die beiden Eigenwerte von A , die ja unimodular sind, zueinander konjugiert.

$\mathbb{T} = \left\{ \begin{pmatrix} \lambda & 0 \\ 0 & \bar{\lambda} \end{pmatrix} \mid \lambda \in S^1 \right\}$ ist eine zu S^1 isomorphe Untergruppe von $SU(2)$.

Auch $SO(2)$ ist eine zu S^1 isomorphe Untergruppe von $SU(2)$.

Das Berechnen der Normalform von $A \in SO(2)$ nach dem in Beispiel 20.5 beschriebenen Verfahren führt hier zu einem konkreten Ergebnis: Ist $A = \begin{pmatrix} a & -b \\ b & a \end{pmatrix}$, $a^2 + b^2 = 1$, so sind $\lambda = a + ib$ und $\bar{\lambda} = a - ib$ Eigenwerte von A und man erhält von A unabhängige Eigenvektoren $v_1 = \begin{pmatrix} i \\ 1 \end{pmatrix}$ zum Eigenwert λ (denn $Av_1 = \begin{pmatrix} ai - b \\ \lambda \end{pmatrix} = \lambda \begin{pmatrix} i \\ 1 \end{pmatrix} = \lambda v_1$)

und $v_2 = \begin{pmatrix} 1 \\ i \end{pmatrix}$ zum Eigenwert $\bar{\lambda}$. Also gilt $S = \frac{-1}{\sqrt{2}} \begin{pmatrix} i & 1 \\ 1 & i \end{pmatrix} \in U(2)$ und

$$S^{-1}AS = \begin{pmatrix} \lambda & 0 \\ 0 & \bar{\lambda} \end{pmatrix}.$$

Damit induziert die unitäre Matrix S einen Gruppenisomorphismus

$$\Phi : SO(2) \rightarrow \mathbb{T}, \quad \Phi(A) = S^{-1}AS.$$

Die Matrix S liegt in $SU(2)$, denn

$$\det S = 1.$$

Man sagt die Untergruppen \mathbb{T} und $SO(2)$ von $SU(2)$ sind **konjugiert** in $SU(2)$.

$$\mathbb{T} = S^{-1} \cdot SO(2) \cdot S.$$

Übungen

- Finden Sie eine Matrix $S \in SU(2)$, so dass $S^{-1} \begin{pmatrix} 1 & i \\ -1 & 1 \end{pmatrix} S$ eine Diagonalmatrix ist.
- Zeigen Sie:

$$A = \frac{1}{1-\omega} \begin{pmatrix} \omega & \omega^2 & \omega^3 \\ \omega^2 & \omega^4 & \omega^6 \\ \omega^3 & \omega^6 & \omega^9 \end{pmatrix} \quad \text{mit } \omega = \frac{1}{2}(-1 + \sqrt{3}i)$$

ist unitär. Berechnen Sie $\det A$.

- Es sei $A = \begin{pmatrix} a & -\bar{b} \\ b & \bar{a} \end{pmatrix} \in SU(2)$, $a, b \in \mathbb{C}$, $|a|^2 + |b|^2 = 1$. Zeigen Sie:
 - Es gibt Diagonalmatrizen $B_1, B_2 \in \mathbb{T}$ (vgl. Beispiel 20.6) und eine Drehmatrix $D \in SO(2)$ mit einem Drehwinkel φ , $0 \leq \varphi \leq \frac{\pi}{2}$, so dass

$$A = B_1 D B_2 \tag{*}$$
 - Sind a, b beide ungleich Null, so ist die Darstellung (*) eindeutig bis auf eine Änderung $(B_1, B_2) \rightarrow (-B_1, -B_2)$.

- Es sei $\mathbb{H} = \left\{ \begin{pmatrix} a & b \\ -\bar{b} & \bar{a} \end{pmatrix} \mid a, b \in \mathbb{C} \right\}$,

$$\mathbf{1} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \mathbf{i} = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}, \quad \mathbf{j} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \quad \mathbf{k} = \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix}.$$

Zeigen Sie:

- (a) \mathbb{H} ist eine nicht-kommutative \mathbb{R} -Unteralgebra von $M_2(\mathbb{C})$, aber keine \mathbb{C} -Unteralgebra.
- (b) \mathbb{H} wird als \mathbb{R} -Vektorraum von $\mathbf{1}, \mathbf{i}, \mathbf{j}, \mathbf{k}$ erzeugt.
- (c) $\mathbb{H}^* = \mathbb{H} \setminus \{0\} \subseteq Gl_2(\mathbb{C})$,
 $\mathbf{H} = \{\pm\mathbf{1}, \pm\mathbf{i}, \pm\mathbf{j}, \pm\mathbf{k}\}$ ist eine Untergruppe von \mathbb{H}^* . Stellen Sie die Multiplikationstafel von \mathbf{H} auf.
- (d) Für $q = a_1 + a_2\mathbf{i} + a_3\mathbf{j} + a_4\mathbf{k}$ sei $\bar{q} = a_1 - a_2\mathbf{i} - a_3\mathbf{j} - a_4\mathbf{k}$. Es gilt $Sp(1) := \{q \in \mathbb{H} \mid q\bar{q} = \mathbf{1}\}$ ist eine zu $SU(2)$ isomorphe Gruppe.

\mathbb{H} ist der **Schiefkörper** der Quaternionen. Der Buchstabe H soll an den Entdecker der Quaternionen, W.R. Hamilton erinnern.

21 Selbstadjungierte Operatoren

Es sei V ein n -dimensionaler unitärer Raum. Wie immer werde das Skalarprodukt von $u, v \in V$ mit $\langle u, v \rangle$ bezeichnet.

Es sei $F : V \rightarrow V$ ein Endomorphismus. Ein besonderer Fall liegt vor, wenn es eine ON -Basis (v_1, \dots, v_n) von V und reelle Zahlen $\lambda_1, \dots, \lambda_n$ gibt, so dass

$$F(v_i) = \lambda_i v_i \text{ für } i = 1, \dots, n.$$

Wir wollen sehen, dass dann für alle $v, w \in V$ gilt

$$\langle F(v), w \rangle = \langle v, F(w) \rangle. \quad (54)$$

Das sieht man folgendermaßen:

Es sei $v = \sum_{i=1}^n x_i v_i$, $w = \sum_{j=1}^n y_j v_j$ mit $x_i, y_j \in \mathbb{C}$.

Dann gilt

$$\begin{aligned} \langle F(v), w \rangle &= \left\langle \sum_{i=1}^n x_i F(v_i), \sum_{j=1}^n y_j v_j \right\rangle = \sum_{i,j=1}^n x_i \bar{y}_j \langle F(v_i), v_j \rangle \\ &= \sum_{i,j=1}^n \lambda_i x_i \bar{y}_j \langle v_i, v_j \rangle = \sum_{i=1}^n \lambda_i x_i \bar{y}_i \end{aligned}$$

und genauso folgt

$$\langle v, F(w) \rangle = \sum_{i=1}^n \bar{\lambda}_i x_i \bar{y}_i.$$

Da nun aber λ_i als reell vorausgesetzt ist, gilt $\bar{\lambda}_i = \lambda_i$ und somit die Formel (54). F ist selbstadjungiert nach der folgenden

Definition 21.1 $F \in \text{End}(V)$ heißt **selbstadjungiert**, wenn $\langle F(v), w \rangle = \langle v, F(w) \rangle$ für alle $v, w \in V$ gilt.

Lemma und Definition 21.2 Zu jedem Endomorphismus $F \in \text{End}(V)$ gibt es genau einen Endomorphismus $F^* \in \text{End}(V)$ mit

$$\langle v, F(w) \rangle = \langle F^*(v), w \rangle$$

für alle $v, w \in V$.

F^* heißt der zu F **adjungierte Operator**.

(Das Wort "linearer Operator auf V " wird synonym zu "Endomorphismus von V " verwendet.)

Es gilt auch

$$\langle F(v), w \rangle = \langle v, F^*(w) \rangle$$

für alle $v, w \in V$, mit anderen Worten: $(F^*)^* = F$.

Ist $\mathcal{B} = (v_1, \dots, v_n)$ eine ON -Basis von V und ist A die Matrix von F bezüglich \mathcal{B} , so ist $A^* := \bar{A}^t$ die Matrix von F^* bezüglich \mathcal{B} .

F ist genau dann selbstadjungiert, wenn $F = F^*$ gilt, d.h. wenn die Matrix A von F bezüglich einer ON -Basis von V **hermitesch** ist, also $A = \bar{A}^t$ gilt.

Beweis: Die Eindeutigkeit von F^* ist klar, weil zwei Vektoren $v_1, v_2 \in V$ genau dann gleich sind, wenn $\langle v_1, w \rangle = \langle v_2, w \rangle$ für alle $w \in V$ gilt.

Zur Existenz von F^* : Es sei $\mathcal{B} = (v_1, \dots, v_n)$ eine ON-Basis von V . Es sei $A = (a_{ij})$ die Matrix von F bezüglich \mathcal{B} , also

$$F(v_j) = \sum_{i=1}^n a_{ij} v_i \text{ für } j = 1, \dots, n.$$

Man definiere $F^* \in \text{End}(V)$ durch

$$F^*(v_i) := \sum_{j=1}^n \bar{a}_{ij} v_j.$$

Dann gilt für $v = \sum x_i v_i$, $w = \sum y_j v_j$:

$$\begin{aligned} \langle v, F(w) \rangle &= \sum_{i,j} x_i \bar{y}_j \langle v_i, F(v_j) \rangle = \sum_{i,j} x_i \bar{y}_j \left\langle v_i, \sum_k a_{kj} v_k \right\rangle \\ &= \sum_{i,j} x_i \bar{y}_j \bar{a}_{ij} \end{aligned}$$

und

$$\begin{aligned} \langle F^*(v), w \rangle &= \sum_{i,j} x_i \bar{y}_j \langle F^*(v_i), v_j \rangle = \sum_{i,j} x_i \bar{y}_j \left\langle \sum_k \bar{a}_{ik} v_k, v_j \right\rangle \\ &= \sum_{i,j} x_i \bar{y}_j \bar{a}_{ij}, \end{aligned}$$

also $\langle F^*(v), w \rangle = \langle v, F(w) \rangle$. Damit ist die Existenz von F^* bewiesen.

Nach Konstruktion ist \bar{A}^t die Matrix von F^* bezüglich \mathcal{B} .

Da $(A^*)^* = A$ gilt, ist A die Matrix von $(F^*)^*$ bezüglich \mathcal{B} , also gilt $F = (F^*)^*$. □

Bemerkung 21.3 Ist V ein n -dimensionaler unitärer Vektorraum, so induziert das Skalarprodukt \langle, \rangle auf V einen kanonischen \mathbb{C} -Vektorraumisomorphismus

$$\text{End}(V) \rightarrow \text{Sesq}(V) := \{h : V \times V \rightarrow \mathbb{C} \mid h \text{ sesquilinear}\}$$

und zwar wird $F \in \text{End}(V)$ die Sesquilinearform h_F mit $h_F(v, w) := \langle F(v), w \rangle$ zugeordnet. Man sieht leicht ein, dass $F \mapsto h_F$ linear ist:

$$\begin{aligned} \text{(a)} \quad h_{F+G}(v, w) &= \langle (F+G)(v), w \rangle = \langle F(v) + G(v), w \rangle = \\ &= \langle F(v), w \rangle + \langle G(v), w \rangle \\ &= h_F(v, w) + h_G(v, w) \\ &= (h_F + h_G)(v, w), \end{aligned}$$

$$\begin{aligned} \text{(b)} \quad h_{\lambda F}(v, w) &= \langle \lambda F(v), w \rangle = \lambda \langle F(v), w \rangle = \\ &= \lambda h_F(v, w). \end{aligned}$$

Weiter gilt $\langle F(v), w \rangle = 0$ für alle $v, w \in V$, genau dann, wenn $F(v) = 0$ für alle $v \in V$, wenn also $F = 0$. Damit ist $F \mapsto h_F$ injektiv. Aus Dimensionsgründen ist damit

$$\text{End}(V) \rightarrow \text{Sesq}(V), \quad F \mapsto h_F$$

ein Isomorphismus.

Bei dieser Zuordnung gilt:

$$F = F^* \iff h_F \text{ ist hermitesche Form.}$$

Ist nämlich $F = F^*$, F also selbstadjungiert, so ist

$$\begin{aligned} h_F(v, w) &= \langle F(v), w \rangle = \langle v, F(w) \rangle = \overline{\langle F(w), v \rangle} \\ &= \overline{h_F(w, v)}, \end{aligned}$$

d.h. h_F ist hermitesch.

Ist umgekehrt h_F hermitesch, so folgt

$$\langle F(v), w \rangle = h_F(v, w) = \overline{h_F(w, v)} = \overline{\langle F(w), v \rangle} = \langle v, F(w) \rangle$$

für alle $v, w \in V$, d.h. F ist selbstadjungiert.

Man beachte, dass die Menge $\text{Herm}(V)$ der hermiteschen Formen $h : V \times V \rightarrow \mathbb{C}$ ein *reeller* aber kein komplexer Untervektorraum von $\text{Sesq}(V)$ ist.

Die Menge der selbstadjungierten Operatoren $F : V \rightarrow V$ ist ein zu $\text{Herm}(V)$ isomorpher reeller Untervektorraum von $\text{End}(V)$.

Man sieht leicht, dass $\dim_{\mathbb{R}} \text{Herm}(V) = n^2$.

Ist $H = (h_{ij})$ die Matrix von h_F bezüglich einer Basis $\mathcal{B} = (v_1, \dots, v_n)$, so gilt (vgl. (19.3.a))

$$h_{ij} = h_F(v_i, v_j) = \overline{h_F(v_j, v_i)} = \overline{h_{ji}},$$

also $H = \overline{H}^t$.

Für die Matrix $A = (a_{ij})$ von F bezüglich \mathcal{B} gilt hingegen

$$\begin{aligned} h_{ij} &= h_F(v_i, v_j) = \langle F(v_i), v_j \rangle = \sum_{k=1}^n a_{ki} \langle v_k, v_j \rangle \\ &= \sum_{k=1}^n a_{ki} g_{kj}, \end{aligned}$$

wobei $G = (g_{ij})$ die Gramsche Matrix von \mathcal{B} ist, $g_{ij} = \langle v_i, v_j \rangle$. Es gilt also

$$H = G^t A = \overline{G} A$$

und aus $H = \overline{H}^t$ folgt $G^t A = \overline{A}^t \overline{G}$, also

$$A = (G^t)^{-1} \overline{A}^t G^t.$$

Nur wenn \mathcal{B} eine *ON*-Basis ist kann man schließen, dass die Matrix A von F bezüglich \mathcal{B} hermitesch ist.

Beispiel 21.4 Interessant ist der Fall $n = 2$.

Sei (e_1, e_2) eine *ON*-Basis V . Die selbstadjungierten Operatoren $F : V \rightarrow V$ werden bezüglich dieser Basis durch hermitesche 2×2 -Matrizen

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

mit $a, c \in \mathbb{R}$, $b \in \mathbb{C}$ beschrieben.

Der reelle Unterraum \mathcal{E} aller selbstadjungierten Operatoren $F : V \rightarrow V$ mit $\text{Spur}(F) = 0$ ist dreidimensional, wird von den Matrizen

$$A = \begin{pmatrix} a & \bar{b} \\ b & -a \end{pmatrix}, a \in \mathbb{R}, b \in \mathbb{C}$$

dargestellt und hat die Basis $(\sigma_1, \sigma_2, \sigma_3)$ bestehend aus den Pauli-Matrizen

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

Man kann sich leicht überlegen, dass durch

$$(F, G) := \frac{1}{2} \text{Spur}(F \circ G)$$

ein euklidisches Skalarprodukt auf \mathcal{E} definiert ist, dessen zugehörige Norm $\|F\|$ durch die Formel $\|F\| = \sqrt{\frac{1}{2} \text{Spur}(A^2)} = \sqrt{a^2 + b\bar{b}} = \sqrt{\det F}$ gegeben ist, wobei $A = \begin{pmatrix} a & \bar{b} \\ b & -a \end{pmatrix}$ die Matrix von F bezüglich der *ON*-Basis (e_1, e_2) ist (Übung).

Wir kommen nun zu dem zentralen Satz über selbstadjungierte Operatoren, dem sogenannten Spektralsatz, der zeigt, dass selbstadjungierte Operatoren sehr einfach gebaut sind.

Satz 21.5 Es sei V ein n -dimensionaler unitärer Vektorraum. Dann gilt

- (a) Ein Operator $F \in \text{End}(V)$ ist genau dann selbstadjungiert, wenn alle Eigenwerte von F reell sind und wenn V eine *ON*-Basis aus Eigenvektoren von F besitzt.
- (b) Eigenvektoren zu verschiedenen Eigenwerten eines selbstadjungierten Operators sind orthogonal.

Beweis zu (a): Wir haben nur die Richtung “ \Rightarrow ” zu beweisen: Es sei $F \in \text{End}(V)$ selbstadjungiert. Ist $\lambda \in \mathbb{C}$ ein Eigenwert von F und $v \in V \setminus \{0\}$ ein Eigenvektor von F zum Eigenwert λ , so gilt

$$\lambda \langle v, v \rangle = \langle \lambda v, v \rangle = \langle F(v), v \rangle = \langle v, F(v) \rangle = \langle v, \lambda v \rangle = \bar{\lambda} \langle v, v \rangle.$$

Da $\langle v, v \rangle \neq 0$ ist, folgt $\lambda = \bar{\lambda}$, d.h. λ ist reell.

Durch Induktion nach n zeigen wir jetzt, dass V eine ON -Basis aus Eigenvektoren von F besitzt.

Für $n = 1$ ist dies offensichtlich.

Es sei $n \geq 2$. Induktionsschluss $n - 1 \rightarrow n$: Wir wählen einen Eigenvektor $v_1 \in V$ zu einem Eigenwert λ von F . Es sei $\|v_1\| = 1$. Wir setzen $U = (\mathbb{C}v_1)^\perp$. Dann ist U $(n - 1)$ -dimensional und F -invariant.

Letzteres sieht man so: Sei $w \in U$, also $\langle v_1, w \rangle = 0$ nach Definition von U . Es folgt

$$\begin{aligned} 0 &= \lambda \langle v_1, w \rangle = \langle \lambda v_1, w \rangle = \langle F(v_1), w \rangle = \\ &= \langle v_1, F(w) \rangle \end{aligned}$$

und somit ist $F(w) \in U$, U also F -invariant. F induziert einen selbstadjungierten Operator $F|U \in \text{End}(U)$. Nach Induktionsvoraussetzung gibt es eine ON -Basis (v_2, \dots, v_n) von U aus Eigenvektoren von $F|U$. Offensichtlich ist dann (v_1, \dots, v_n) eine ON -Basis von V aus Eigenvektoren. \square

zu (b): Sind $\lambda, \mu \in \mathbb{R}$, $\lambda \neq \mu$, $v, w \in V \setminus \{0\}$ mit $F(v) = \lambda v$, $F(w) = \mu w$, so folgt

$$\begin{aligned} \lambda \langle v, w \rangle &= \langle \lambda v, w \rangle = \langle F(v), w \rangle = \langle v, F(w) \rangle \\ &= \langle v, \mu w \rangle = \mu \langle v, w \rangle, \end{aligned}$$

also $(\lambda - \mu)\langle v, w \rangle = 0$. Da $\lambda - \mu \neq 0$, folgt $\langle v, w \rangle = 0$. \square

Korollar 21.6 Ist $A \in M_n(\mathbb{C})$ eine hermitesche Matrix ($A = \overline{A}^t$), so ergibt es eine unitäre Matrix $S \in U(n)$, so dass

$$S^{-1}AS = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix} \text{ mit } \lambda_i \in \mathbb{R}$$

Beweis: Man fasse $A : \mathbb{C}^n \rightarrow \mathbb{C}^n$ als selbstadjungierten Endomorphismus auf. Nach Satz 21.5 (a) gibt es eine ON -Basis (v_1, \dots, v_n) von \mathbb{C}^n , so dass $Av_i = \lambda_i v_i$ mit $\lambda_i \in \mathbb{R}$. Es sei S die Matrix mit den Spalten v_1, \dots, v_n . Dann ist $S \in U(n)$ und

$$AS = S \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}.$$

\square

Bemerkung 21.7 Es sei V ein n -dimensionaler unitärer Vektorraum und $F \in \text{End}(V)$ selbstadjungiert. $\lambda_1, \dots, \lambda_m$ seien die verschiedenen Eigenwerte von F , λ_i trete mit der Vielfachheit $n_i \geq 1$ auf.

Dann ist $E_i = E(\lambda_i) = \ker(F - \lambda_i \text{id}_V)$ ein n_i -dimensionaler Unterraum von V .

Es sei $P_{E_i} \in \text{End}(V)$ der orthogonale Projektor

$$P_{E_i}(v + w) = v, \text{ wenn } v \in E_i, w \in E_i^\perp.$$

Nach dem Spektralsatz gilt dann

$$\begin{aligned} F(v_1 + \dots + v_m) &= F(v_1) + \dots + F(v_m) = \\ &= \lambda_1 v_1 + \dots + \lambda_m v_m = \\ &= \lambda_1 P_{E_1}(v_1 + \dots + v_m) + \dots + \lambda_m P_{E_m}(v_1 + \dots + v_m) \end{aligned}$$

für $v = v_1 + \dots, v_m, v_i \in E_i$, kurz:

$$F = \sum_{i=1}^m \lambda_i P_{E_i} \quad (55)$$

Diese Darstellung von F heißt die **Spektralzerlegung** von F .

Der Spektralsatz gilt allgemeiner für gewisse Operatoren auf einem Hilbertraum \mathcal{H} . Um den Begriff des Hilbertraumes zu erklären bedarf es einiger Hilfsmittel aus der Analysis, insbesondere des Begriffs der Konvergenz.

Ein Hilbertraum ist ein (unendlich-dimensionaler) \mathbb{C} -Vektorraum \mathcal{H} mit einem positiv definiten unitären Skalarprodukt $\langle \cdot, \cdot \rangle$, so dass jede Cauchy-Folge $(\varphi_n)_{n \in \mathbb{N}}$ von Elementen $\varphi_n \in \mathcal{H}$ in \mathcal{H} konvergiert. (siehe: Forster: Analysis I). Die Hilbertraumtheorie ist von fundamentaler Bedeutung in der Quantenmechanik.

Man lernt diese Theorie in Vorlesungen über Funktionalanalysis (ab dem 3. oder 4. Semester) oder im Selbststudium etwa mit Riesz/Sz.-Nagy: Vorlesungen über Funktionalanalysis.

Beispiel 21.8 Wir greifen noch einmal das Beispiel 19.8 auf. Die komplexwertigen Funktionen

$$\varphi_n(x) = \frac{1}{\sqrt{2\pi}} e^{inx}, \quad |n| < N, x \in [0, 2\pi]$$

spannen einen $(2N + 1)$ -dimensionalen Untervektorraum V_N von $\text{Abb}([0, 2\pi], \mathbb{C})$ auf. Mit

$$\langle \varphi, \psi \rangle = \int_0^{2\pi} \varphi(x) \overline{\psi(x)} dx$$

ist V_N ein unitärer Raum und, wie wir gesehen haben, ist $\{\varphi_n \mid |n| \leq N\}$ eine Orthonormalbasis von V_N . Es gilt

$$\overline{\varphi_n} = \varphi_{-n}$$

und

$$\varphi_0 = \frac{1}{\sqrt{2\pi}}.$$

Nun ist $\frac{d}{dx} e^{inx} = ine^{inx}$ und somit $\frac{d^2}{dx^2} e^{inx} = i^2 n^2 e^{inx} = -n^2 e^{inx}$, also ist der sogenannte (eindimensionale) *Laplace-Operator* $\Delta := \frac{d^2}{dx^2}$ ein linearer Operator auf V_N .

$$\Delta : V_N \rightarrow V_N.$$

Mit Hilfe partieller Integration zeigt man, dass Δ selbstadjungiert ist:

$$\begin{aligned}\langle \Delta\varphi, \psi \rangle &= \int_0^{2\pi} \frac{d}{dx} \left(\frac{d}{dx} \varphi(x) \right) \overline{\psi(x)} dx = \left[\frac{d}{dx} \varphi(x) \cdot \overline{\psi(x)} \right]_0^{2\pi} - \int_0^{2\pi} \frac{d}{dx} \varphi(x) \frac{d}{dx} \overline{\psi(x)} dx \\ &= - \int_0^{2\pi} \frac{d}{dx} \varphi(x) \overline{\frac{d}{dx} \psi(x)} dx = \int_0^{2\pi} \varphi(x) \overline{\Delta\psi(x)} dx = \langle \varphi, \Delta\psi \rangle.\end{aligned}$$

Es gilt $\Delta\varphi_n = -n^2\varphi_n$, also ist φ_n eine "Eigenvektor" von Δ mit Eigenwert $-n^2$.

Wir nennen φ_n in diesem Zusammenhang eine **Eigenfunktion** von Δ .

Damit ist $(\varphi_{-N}, \dots, \varphi_0, \varphi_1, \dots, \varphi_N)$ eine Orthonormalbasis auf Eigenfunktionen von Δ . Die zugehörigen Eigenräume der Eigenwerte $0, -1, -4, \dots, -N^2$ sind

$$\begin{aligned}E_0 &= \mathbb{C}\varphi_0 \text{ konstante Funktionen} \\ E_{-n^2} &= \ker(\Delta + n^2 id_{V_N}) = \mathbb{C}\varphi_n \oplus \mathbb{C}\overline{\varphi_n}\end{aligned}$$

Damit haben wir die Spektralzerlegung des Laplace-Operators auf V_N gefunden.

$$\Delta = -P_{E_{-1}} - 4P_{E_{-4}} - 9P_{E_{-9}} - \dots - N^2 P_{E_{-N^2}}.$$

Der Laplace-Operator (in mehreren Veränderlichen) spielt eine sehr wichtige Rolle in der Analysis, Differentialgeometrie und der theoretischen Physik.

Übungen

- Es sei V ein unitärer Vektorraum von $F : V \rightarrow V$ ein selbstadjungierter Operator. F heißt nichtnegativ ($F \geq 0$), wenn $\langle F(v), v \rangle \geq 0$ für alle $v \in V$ gilt.
 - Beweisen Sie: $F \geq 0 \iff$ Alle Eigenwerte von F sind ≥ 0 .
 - Beweisen Sie: Sind $F, G \in \text{End}(V)$ selbstadjungiert und gilt $FG = GF$, so gilt:

$$F \geq G \text{ und } G \geq 0 \Rightarrow FG \geq 0.$$
 - Ist $F \geq 0$, so gibt es genau einen selbstadjungierten Operator $G \in \text{End}(V)$, $G \geq 0$, so dass $F = G^2$.
- Sei V ein unitärer Vektorraum und $F \in \text{End}(V)$. Dann ist $G := F^* \circ F$ selbstadjungiert und es gilt $G \geq 0$. Es gilt $G > 0$ (d.h. $\langle G(v), v \rangle > 0$ für alle $v \in V \setminus \{0\}$) genau dann, wenn F ein Isomorphismus ist.
- Es sei $n \in \mathbb{N}$ fest gewählt und V sei der reelle Vektorraum der Polynome $f \in \mathbb{R}[x]$ mit $\text{Grad}(f) \leq n$. Es sei

$$\langle f, g \rangle := \int_{-1}^1 f(x)g(x)dx$$

Dann ist \langle, \rangle ein Skalarprodukt auf V .

Zeigen Sie: $F = (x^2 - 1)\frac{d^2}{dx^2} + 2x\frac{d}{dx}$ ist ein selbstadjungierter Operator auf V .

Hinweis: Berechnen Sie $F(P_k)$, wobei P_k die Legendre-Polynome sind:

$$P_k(x) = \frac{1}{2^k k!} \frac{d^k}{dx^k} (x^2 - 1)^k \text{ für } k \geq 1, P_0 = 1.$$

4. Formulieren und beweisen Sie den Spektralsatz für euklidische Vektorräume.

5. Es sei

$$A = \begin{pmatrix} \frac{5}{4} & \frac{1}{2\sqrt{2}} & -\frac{3}{4} \\ \frac{1}{2\sqrt{2}} & \frac{1}{2} & \frac{1}{2\sqrt{2}} \\ -\frac{3}{4} & \frac{1}{2\sqrt{2}} & \frac{5}{4} \end{pmatrix}$$

Bestimmen Sie eine Drehmatrix $S \in SO(3)$, so dass $S^{-1}AS$ eine Diagonalmatrix ist.

6. Es sei V ein unitärer Vektorraum und $P \in \text{End}(V)$ sei selbstadjungiert und es gelte $P^2 = P$.

Zeigen Sie: Es gibt eine orthogonale Zerlegung $V = U \oplus W$ von V , so dass P der orthogonale Projektor von V auf U ist.

7. Beweisen Sie die Behauptung aus Beispiel 21.4.

Zeigen Sie weiter: Für jede Matrix $S \in SU(2)$ ist die Abbildung $A \mapsto SAS^{-1}$ eine orthogonale Abbildung $\Phi_S: \mathcal{E} \rightarrow \mathcal{E}$ bezüglich des in 21.4 definierten Skalarproduktes auf \mathcal{E} .

Zeigen Sie weiter, dass $S \mapsto \Phi_S$ einen surjektiven Gruppenhomomorphismus

$$\Phi: SU(2) \rightarrow SO(3) = SO(3)$$

definiert. Bestimmen Sie den Kern $\{S \mid \Phi_S = id_{\mathcal{E}}\}$.

8. Eine Matrix $S \in M_n(\mathbb{C})$ heißt schieferhermitesch, wenn $\overline{S}^t = -S$ gilt. Zeigen Sie, dass $S \mapsto iS$ einen \mathbb{R} -Vektorraumisomorphismus vom Raum der hermiteschen auf den Raum der schieferhermiteschen Matrizen liefert.

Zeigen Sie weiter, dass $im\mathbb{H} := \{a\mathbf{i} + b\mathbf{j} + c\mathbf{k} \mid a, b, c \in \mathbb{R}\}$ der Raum der schieferhermiteschen komplexen 2×2 -Matrizen mit Spur Null ist.

Wie muss man das Skalarprodukt auf \mathbb{H} einführen, damit die Abbildung $A \mapsto iA$ eine Isometrie $\mathcal{E} \rightarrow im\mathbb{H}$ definiert?

Bilden die Quaternionen $\mathbf{i}, \mathbf{j}, \mathbf{k}$ eine Orthonormalbasis von $im\mathbb{H}$?

(\mathbb{H} ist wie in Übung 20.4 definiert.)

22 Quadratische Formen

In diesem letzten Abschnitt der Vorlesung wird der Begriff des euklidischen Vektorraum weiter verallgemeinert. Diese Verallgemeinerung ist von Bedeutung in der Zahlentheorie aber auch in der Analysis, der Relativitätstheorie und nicht zuletzt in der algebraischen Topologie.

Im folgenden sei K ein Körper der Charakteristik $\neq 2$.

Definition 22.1 Ein **orthogonaler Raum** (V, g) über K besteht aus einem endlich-dimensionalen K -Vektorraum V und einer symmetrischen Bilinearform

$$g : V \times V \rightarrow K.$$

Ist $\mathcal{B} = (v_1, \dots, v_n)$ eine Basis von V , so ist g durch die symmetrische Matrix $G = (g(v_i, v_j)) \in M_n(K)$ bestimmt: Ist $v = \sum_{i=1}^n x_i v_i$, $w = \sum_{i=1}^n y_i v_i$, so gilt

$$g(v, w) = \sum_{i,j=1}^n x_i g(v_i, v_j) y_j.$$

Beispiel 22.2 Jede symmetrische Matrix $G \in M_n(\mathbb{R})$ definiert eine orthogonale Geometrie auf \mathbb{R}^n . Ist G positiv definit, so ist die Geometrie **euklidisch**. Es kommen aber auch andere Geometrien vor. In der speziellen Relativitätstheorie ist die Geometrie (\mathbb{R}^4, g) , $g(x, y) = +x_1 y_1 + x_2 y_2 + x_3 y_3 - x_4 y_4$ von fundamentaler Bedeutung (der Minkowski-Raum).

Definition 22.3 Es sei (V, g) ein orthogonaler Raum über K .

- (a) Für $v, w \in V$ definiert man (wie im euklidischen Fall): v und w sind orthogonal (bezüglich g), in Zeichen $v \perp w$ genau dann, wenn $g(v, w) = 0$. \perp ist eine symmetrische Relation auf V , weil g symmetrisch ist.
- (b) Sind $U, W \subset V$ K -Untervektorräume, so heißen U, W **orthogonal** ($U \perp W$) wenn $u \perp w$ für alle $u \in U$, $w \in W$ gilt.

$$U^\perp := \{v \in V \mid g(u, v) = 0 \text{ für alle } u \in U\}$$

heißt der **Orthogonalraum** zu U .

Wenn g nicht positiv definit ist, ist die Relation \perp andersartig als die gewohnte anschauliche Bedeutung von \perp im euklidischen Fall. Das führt das folgende Beispiel drastisch vor Augen.

Beispiel 22.4 Auf $V = \mathbb{R}^2$ betrachte man die indefinite Form $g(x, y) = x_1 y_1 - x_2 y_2$. Für $x = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ gilt dann offensichtlich $g(x, x) = 0$, d.h. $x \perp x$, obwohl $x \neq 0$.

Ist U der von x aufgespannte Unterraum, so ist $U^\perp = \{y = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \in \mathbb{R}^2 \mid g(x, y) = 0\} = U$. Für $x = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$, $U = \mathbb{R}x$ gilt dagegen $U^\perp = \{\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \mid y_1 = 0\} = \mathbb{R}y$, wobei $y = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$.

Hier ist $U \cap U^\perp = 0$ und $V = U \oplus U^\perp$.

Also: Der Orthogonalraum eines Unterraums U kann, muss aber nicht notwendig komplementär zu U sein.

Definition 22.5 Ist (V, g) ein orthogonaler Raum, so definiert g eine lineare Abbildung

$$\tilde{g} : V \rightarrow V^*$$

von V in seinen Dualraum V^* und zwar in kanonischer Weise:

$$\tilde{g}(v)(w) = g(v, w).$$

Es gilt

$$V^\perp = \{v \in V \mid g(v, w) = 0 \text{ für alle } w \in V\} = \ker \tilde{g}$$

g heißt **nicht-entartet** $\iff \ker \tilde{g} = 0$ ($\iff \tilde{g}$ ist ein Isomorphismus).

Der Rang von \tilde{g} heißt auch der **Rang** von g . g ist genau dann nicht-entartet, wenn die Matrix G von g bzgl. einer Basis von V invertierbar ist.

Definition 22.6 Ein **Isomorphismus** zwischen orthogonalen Räumen

$(V, g), (V', g')$ ist ein Vektorraumisomorphismus $F : V \rightarrow V'$, der die Bilinearformen respektiert, d.h. für den $g'(F(v), F(w)) = g(v, w)$ für alle $v, w \in V$ gilt. Man nennt diese Isomorphismen auch **Isometrien**.

$$O(V, g) = \{F : V \rightarrow V \mid F \text{ Isometrie}\}$$

ist eine Untergruppe von $GL(V)$ und heißt die **orthogonale Gruppe** von g .

Das Klassifikationsproblem für orthogonale Räume besteht nun darin, die orthogonalen Räume (V, g) über K bis auf Isomorphie zu bestimmen. Im Fall $K = \mathbb{C}$ und auch im Fall $K = \mathbb{R}$ gibt es eine einfache Lösung.

Zunächst betrachten wir den Fall $n = \dim V = 1$. Dann ist eine Bilinearform $g : V \times V \rightarrow K$ vollständig durch $a = g(v, v)$ bestimmt, wobei $v \in V \setminus 0$ ein fest gewählter Basisvektor sei. Es gilt

$$g(\lambda v, \mu v) = \lambda \mu a \text{ für alle } \lambda \mu \in K.$$

Ist nun (V', g') ein weiterer eindimensionaler orthogonaler Raum über K , $v' \in V' \setminus 0$ und $a' = g'(v', v')$, so gilt

$$(V, g) \cong (V', g') \iff \exists \text{ Isometrie } F : (V, g) \rightarrow (V', g').$$

Solch eine Isometrie ist durch $b \in K^\times$ mit $F(v) = bv'$ bestimmt. F respektiert die Bilinearformen g, g' genau dann, wenn $a = b^2 a'$, denn

$$a = g(v, v) \text{ und } g'(F(v), F(v)) = g'(bv', bv') = b^2 g'(v', v') = b^2 a'.$$

Es gilt also

$$(V, g) \cong (V', g') \Rightarrow \exists b \in K^\times : a = b^2 a',$$

d.h. a und a' unterscheiden sich um ein "Quadrat" in K^\times .

Damit haben wir:

Lemma 22.7 Die 1-dimensionalen nicht-entarteten orthogonalen Räume entsprechen genau den Äquivalenzklassen

$$K^{\times 2}a = \{b^2a \mid b \in K^\times\}$$

wobei $a \in K^\times$.

Für $a = 0$ bekommt man den entarteten orthogonalen Raum (V, g) mit $g = 0$. \square

Man kann sich leicht überlegen, dass die Menge

$$K^\times / K^{\times 2} := \{K^{\times 2}a \mid a \in K^\times\}$$

mit der Multiplikation

$$(K^{\times 2}a)(K^{\times 2}b) = K^{\times 2}ab$$

eine abelsche Gruppe ist.

Die Ordnung dieser Gruppe gibt an, wie viele wesentlich verschiedene, d.h. nicht zueinander isomorphe eindimensionale nicht-entartete orthogonale Räume es gibt.

Beispiel 22.8

- (a) $K = \mathbb{C}$. Hier ist $\mathbb{C}^\times = \mathbb{C}^{\times 2}$, denn jede komplexe Zahl $z \in \mathbb{C}^\times$ hat eine Quadratwurzel. Es gibt also bis auf Isomorphie genau einen nicht-entarteten 1-dimensionalen orthogonalen Raum über \mathbb{C} , nämlich \mathbb{C} mit Standardform $g(z, w) = zw$.
- (b) $K = \mathbb{R}$. Hier ist $\mathbb{R}^{\times 2} = \mathbb{R}_+$ und $\mathbb{R}^\times / \mathbb{R}_+$ besteht aus zwei Elementen, der Klasse von 1 und der Klasse von -1. Damit gibt es zwei inäquivalente 1-dimensionale nicht-entartete orthogonale Räume über \mathbb{R} : $(\mathbb{R}, (1))$, $(\mathbb{R}, (-1))$, den **positiven** und den **negativen** orthogonalen Raum.
- (c) Ist $K = \mathbb{F}_p$, p ungerade Primzahl, so kann man sich leicht überlegen, dass es genau $\frac{p-1}{2}$ Quadrate in \mathbb{F}^\times gibt. Die Fasern der surjektiven Abbildung $\mathbb{F}^\times \rightarrow \mathbb{F}_p^{\times 2}$, $x \mapsto x^2$ bestehen nämlich stets aus zwei Elementen.

Damit gibt es über \mathbb{F}_p genau zwei inäquivalente 1-dimensionale nicht-entartete orthogonale Räume.

Definition 22.9 Es sei (V, g) ein orthogonaler Raum über K . Ein Untervektorraum $W \subset V$ heißt

- (a) **nicht-entartet** \iff die Einschränkung von g auf $W \times W$ ist nicht-entartet.
- (b) **isotrop** \iff die Einschränkung von g auf $W \times W$ ist trivial, d.h. identisch Null (also: $W \subset W^\perp$).

Beispiel 22.10 (Minkowski-Raum)

$\mathcal{M} = (\mathbb{R}^4, g)$ mit

$$g(x, y) = x_1y_1 + x_2y_2 + x_3y_3 - x_4y_4$$

heißt der **Minkowski-Raum**. Es gibt hier isotrope Unterräume.

Es sei etwa $W = \mathbb{R}x$, wobei $x := (1, 0, 0, 1)$. Dann ist $g(x, x) = 1 - 1 = 0$, also ist $W \subset W^\perp$. In W^\perp liegen aber auch noch weitere Vektoren, z.B. $(0, 1, 0, 0)$, $(0, 0, 1, 0)$.

Um dieses Beispiel weiter untersuchen zu können, brauchen wir noch etwas mehr Theorie. Wir bemerken lediglich, dass dieser Raum in der speziellen Relativitätstheorie als *Raumzeit* auftritt. Die letzte Koordinate x_4 ist die Größe ct , wobei c die Lichtgeschwindigkeit bezeichnet und t die Zeit. Punkte x des Minkowskiraums -man nennt sie auch *Weltpunkte*- heißen

- *raumartig*, wenn $g(x, x) > 0$,
- *zeitartig*, wenn $g(x, x) < 0$,
- *lichtartig*, wenn $g(x, x) = 0$.

Damit hat es die folgende Bewandnis. Die Weltpunkte, die von 0 aus durch ein Lichtsignal erreicht werden, sind die Punkte auf dem Kegel $g(x, x) = 0, x_4 > 0$. Dies sind die lichtartigen Punkte. Ist $p = tv$ die Bahn eines sich mit konstanter Geschwindigkeit $v \in \mathbb{R}^3$ bewegendes Punktes, so ist $L = \{(tv, ct) \mid t \in \mathbb{R}\}$ die zugehörige Weltlinie, die wegen $\|v\| < c$ mit Ausnahme von 0 nur aus zeitartigen Punkten besteht.

Satz 22.11 Es sei (V, g) ein orthogonaler Raum über K um $U \subset V$ ein Untervektorraum. Dann gilt

- (a) Ist U nicht-entartet, so ist U^\perp komplementär zu U , d.h. es gilt $V = U \oplus U^\perp$.

Man nennt in diesem Fall U^\perp das orthogonale Komplement von U .

- (b) Sind U und U^\perp nicht-entartet, so gilt

$$(U^\perp)^\perp = U.$$

Beweis: (a) Da U nicht-entartet ist, ist $f : U \rightarrow U^*$ mit $f(x)(y) = g(x, y)$ ein Isomorphismus. Wir zeigen, dass $U \cap U^\perp = 0$ gilt. Sei dazu $x \in U \cap U^\perp$. Dann gilt $x \in U$ und $g(x, y) = 0$ für alle $y \in U$, also ist $f(x) = 0$ und somit $x = 0$. Jetzt zeigen wir, dass $U + U^\perp = V$ gilt. Ist $x \in V$ so betrachten wir $\varphi \in U^*$ mit $\varphi(u) = g(u, x)$ für $u \in U$.

Da f bijektiv ist, gibt es ein eindeutig bestimmtes $x' \in U$ so dass $f(x') = \varphi$, d.h. $g(u, x') = g(u, x)$ für alle $u \in U$.

Sei $x'' = x - x'$. Dann ist $g(u, x'') = g(u, x) - g(u, x') = 0$ für alle $u \in U$, also $x'' \in U^\perp$.

Damit haben wir die gewünschte Zerlegung

$$x = x' + x''$$

gefunden.

(b) $U \subset (U^\perp)^\perp$ ist klar, denn ist $x \in U$, so ist $g(x, y) = 0$ für alle $y \in U^\perp$, also $x \in (U^\perp)^\perp$. Nach Teil (a) ist $\dim U = \dim V - \dim U^\perp$. Da U^\perp nicht-entartet sein soll, ist auch

$$\dim U^\perp = \dim V - \dim (U^\perp)^\perp.$$

Es folgt $\dim U = \dim (U^\perp)^\perp$, also $U = (U^\perp)^\perp$. □

Aus diesem Satz folgt nun induktiv der sehr wichtige Zerlegungssatz.

Satz 22.12 Es sei (V, g) ein n -dimensionaler orthogonaler Raum über K ($\text{char } K \neq 2$). Dann gibt es eine Zerlegung

$$V = V_1 \oplus \dots \oplus V_n$$

in paarweise orthogonale 1-dimensionale Untervektorräume V_i von V .

Beweis: Induktion nach n .

Für $n = 1$ ist nichts zu beweisen.

Es sei $n \geq 2$ und die Behauptung für Räume der Dimension $< n$ sei schon bewiesen.

Ist $g = 0$, so ist jede Zerlegung in 1-dimensionale Unterräume erlaubt, weil $g(x, y) = 0 \forall x, y \in V$.

Sei also jetzt $g \neq 0$.

Es gibt dann Vektoren $x, y \in V$, so dass $g(x, y) \neq 0$.

Es folgt für $z = x + y$:

$$g(z, z) = g(x, x) + 2g(x, y) + g(y, y).$$

Ist nun $g(x, x) = 0$ und $g(y, y) = 0$, so ist (wegen $\text{char}(K) \neq 2$) $g(z, z) \neq 0$. Es gibt also in jedem Fall, wenn $g \neq 0$, auch einen Vektor $x \in V$ mit $g(x, x) \neq 0$.

$U := Kx$ ist dann ein 1-dimensionaler nicht-entarteter Unterraum von V . Nach Satz 22.11 folgt

$$U \oplus U^\perp = V$$

und $\dim U^\perp = n - 1$. Nach Induktionsvoraussetzung sind wir fertig. \square

Korollar 22.13 Jeder orthogonale Raum (V, g) über K ($\text{char}(K) \neq 2$) besitzt eine **Orthogonalbasis** d.h. eine Basis (v_1, \dots, v_n) von V mit der Eigenschaft $g(v_i, v_j) = 0$ für $i \neq j$.

Ist g nicht-entartet, so gilt für eine Orthogonalbasis (v_1, \dots, v_n) stets $g(v_i, v_i) \neq 0$ für $i = 1, \dots, n$.

Beweis: Ist $V = V_1 \oplus \dots \oplus V_n$ orthogonale Zerlegung von V in eindimensionale Unterräume, so wähle $v_i \in V_i$, $v_i \neq 0$. Die Matrix von g bzgl. (v_1, \dots, v_n) ist

$$G = \begin{pmatrix} g(v_1, v_1) & & 0 \\ & \ddots & \\ 0 & & g(v_n, v_n) \end{pmatrix}.$$

Ist g nicht-entartet, so muss für alle $i = 1, \dots, n$ $g(v_i, v_i) \neq 0$ gelten. \square

Zusatz: Ist $K = \mathbb{C}$, so kann man eine Orthogonalbasis (v_1, \dots, v_n) so wählen, dass $g(v_i, v_i) = 0$ oder 1.

Ist $K = \mathbb{R}$, so kann man eine Orthogonalbasis so wählen, dass $g(v_i, v_i) = 0, 1$ oder -1 . Dies erfolgt aus Beispiel 22.8(a),(b). \square

Korollar 22.14 ($\text{char}(K) \neq 2$)

Ist $G \in M_n(K)$ symmetrische Matrix, so gibt es eine invertierbare Matrix $S \in GL_n(K)$, so dass

$$S^t G S = \begin{pmatrix} a_1 & & 0 \\ & \ddots & \\ 0 & & a_n \end{pmatrix}$$

mit $a_1, \dots, a_n \in K$.

Ist $K = \mathbb{C}$, so kann man $a_i \in \{0, 1\}$,

Ist $K = \mathbb{R}$, so kann man $a_i \in \{0, 1, -1\}$ wählen. \square

Korollar 22.15 (Klassifikationssatz über \mathbb{C})

Zwei orthogonale Räume (V, g) , (V', g') über \mathbb{C} sind genau dann isomorph, wenn

$$\dim V = \dim V' \text{ und } \operatorname{rg}(g) = \operatorname{rg}(g')$$

gilt. Insbesondere gibt es bis auf Isomorphie genau einen nicht-entarteten orthogonalen Raum der Dimension n , nämlich \mathbb{C}^n mit der Standardform $g(z, w) = \sum_{j=1}^n z_j w_j$ (nicht zu verwechseln mit dem unitären Standardskalarprodukt). \square

Etwas schwieriger ist die Klassifikation über \mathbb{R} .

Satz 22.16 (Sylvesterscher Trägheitssatz)

Es sei (V, g) ein n -dimensionaler orthogonaler Raum über \mathbb{R} . Dann gibt es zwei Zahlen $r_+, r_- \in \mathbb{N}$, so dass für alle orthogonalen Zerlegungen

$$V = V_1 \oplus \dots \oplus V_n$$

in eindimensionale Teilräume gilt:

$$r_+ \text{ ist die Anzahl der positiven Teilräume } V_j$$

und

$$r_- \text{ ist die Anzahl der negativen Teilräume } V_j$$

oder anders formuliert:

Ist (v_1, \dots, v_n) irgendeine Orthogonalbasis von V und $a_i = g(v_i, v_i)$ so ist

$$\begin{aligned} r_+ &= |\{i \mid a_i > 0\}| \text{ und} \\ r_- &= |\{i \mid a_i < 0\}| \end{aligned}$$

Insbesondere ist $r_+ + r_- = \operatorname{rg}(g)$.

Beweis: Da der Rang von g eine Invariante von g ist, müssen wir nur zeigen, dass die Anzahl der positiven Basisvektoren in einer Orthogonalbasis unabhängig von der gewählten Orthogonalbasis ist.

Seien dazu zwei Orthogonalbasen (v_1, \dots, v_n) und (v'_1, \dots, v'_n) gegeben. Es gelte

$$\begin{aligned} g(v_i, v_i) > 0 &\iff 1 \leq i \leq r \\ g(v'_i, v'_i) > 0 &\iff 1 \leq i \leq s \end{aligned}$$

Behauptung: $r = s$

Annahme: $r > s$

Es sei $U = \text{Span}(v_1, \dots, v_r)$,

$$U' = \text{Span}(v'_1, \dots, v'_s), \quad U'' = \text{Span}(v'_{s+1}, \dots, v'_n).$$

Für $w \in U''$ gilt $g(w, w) \leq 0$, denn ist

$$w = \sum_{i>s} a_i v'_i, \quad \text{so ist } g(w, w) = \sum_{i>s} a_i^2 \underbrace{g(v'_i, v'_i)}_{\leq 0} \leq 0.$$

Da $V = U' \oplus U''$, kann man nun v_j folgendermaßen darstellen:

$$v_j = u_j + w_j \quad \text{mit } u_j \in U', w_j \in U''$$

für $j = 1, \dots, r$. Da $\dim U' = s < r$, sind u_1, \dots, u_r linear abhängig. Es gibt also ein $(a_1, \dots, a_r) \in \mathbb{R}^r \setminus \{0\}$ so dass $a_1 u_1 + \dots + a_r u_r = 0$. Es sei $v = a_1 v_1 + \dots + a_r v_r$.

Es folgt $g(v, v) = a_1^2 g(v_1, v_1) + \dots + a_r^2 g(v_r, v_r) > 0$, denn nach Voraussetzung ist $g(v_i, v_i) > 0$ für $1 \leq i \leq r$.

Nun ist aber andererseits

$$v = a_1(u_1 + w_1) + \dots + a_r(u_r + w_r) = a_1 w_1 + \dots + a_r w_r \in U'',$$

also $g(v, v) \leq 0$, wie wir oben bemerkt haben. Damit ist die Annahme zum Widerspruch geführt! Es gilt $r = s$. \square

Definition 22.17 Das Paar (r_+, r_-) heißt die **Signatur** von (V, g) .

Bemerkung 22.18 Es sei V ein euklidischer Vektorraum der Dimension n , $\langle x, y \rangle$ sei das Skalarprodukt von x und y .

V ist ein orthogonaler Raum mit der Signatur $(n, 0)$. Nun ist der Vektorraum der selbstadjungierten Operatoren

$$F : V \rightarrow V$$

d.h. der Endomorphismen $F : V \rightarrow V$ mit $\langle v, F(w) \rangle = \langle F(v), w \rangle$ via $\langle \cdot, \cdot \rangle$ isomorph zu dem Raum $\text{Sym}_2(V)$ der symmetrischen Bilinearformen auf V . Der Isomorphismus ist durch $F \mapsto g_F$ mit $g_F(x, y) := \langle F(x), y \rangle$ gegeben. Nach dem Spektralsatz gibt es eine Orthonormalbasis von V , so dass die Matrix A von F bzgl. dieser Basis eine Diagonalmatrix ist mit den Eigenwerten von F in der Diagonale.

Die Signatur (r_+, r_-) von g_F kann man an den Eigenwerten von F ablesen:

r_+ ist die Anzahl der positiven und
 r_- ist die Anzahl der negativen Eigenwerte von F .

Ist $\mathcal{B} = (v_1, \dots, v_n)$ eine beliebige Orthogonalbasis von V , so ist die Matrix A von F bezüglich \mathcal{B} eine symmetrische Matrix und A ist auch gleichzeitig die Matrix von g_F bezüglich \mathcal{B} (vgl. 21.3).

Nach dem Spektralsatz gibt es eine orthogonale Matrix $S \in O(n)$, so dass

$$S^{-1}AS = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}.$$

Das charakteristische Polynom von A hat dann die Form $p_A = (x - \lambda_1) \dots (x - \lambda_n)$. Nach dem Korollar 22.14 gibt es eine invertierbare Matrix $T \in GL_n(\mathbb{R})$, so dass

$$T^t AT = \left(\begin{array}{cccccccc} 1 & & & & & & & 0 \\ & \ddots & & & & & & \\ & & 1 & & & & & \\ & & & -1 & & & & \\ & & & & \ddots & & & \\ & & & & & -1 & & \\ 0 & & & & & & 0 & \ddots \\ & & & & & & & & 0 \end{array} \right) \left. \begin{array}{l} \vphantom{\begin{pmatrix} 1 \\ \vdots \\ 1 \\ -1 \\ \vdots \\ -1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}} \right\} r_+ \\ \left. \vphantom{\begin{pmatrix} 1 \\ \vdots \\ 1 \\ -1 \\ \vdots \\ -1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}} \right\} r_-$$

Da $S \in O(n)$, gilt auch $S^{-1} = S^t$, also

$$(S^t T)^t \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix} (S^t T) = \begin{pmatrix} 1 & & & & & & & 0 \\ & \ddots & & & & & & \\ & & 1 & & & & & \\ & & & -1 & & & & \\ & & & & \ddots & & & \\ & & & & & -1 & & \\ 0 & & & & & & 0 & \ddots \\ & & & & & & & & 0 \end{pmatrix}.$$

Wir haben zwei verschiedene Äquivalenzrelationen auf dem Vektorraum $S_n(\mathbb{R})$ der symmetrischen $n \times n$ -Matrizen. Die erste bekommt man, wenn man $A \in S_n(\mathbb{R})$ als selbstadjungierten Operator

$$A : \mathbb{R}^n \rightarrow \mathbb{R}^n$$

auffasst. Dann gilt für $A, B \in S_n(\mathbb{R})$:

$$A \sim B \quad (A \text{ ist äquivalent zu } B, \text{ auch } A \text{ ist } \mathbf{\ddot{a}hnlich} \text{ zu } B) \iff B = S^{-1}AS \text{ f\u00fcr ein } S \in GL_n(\mathbb{R}).$$

Fasst man dagegen die Elemente aus $S_n(\mathbb{R})$ als symmetrische Bilinearformen

$$A : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$$

auf, so kommt man zu der Äquivalenzrelation

$$A \cong B \quad (A \text{ ist } \mathbf{kongruent} \text{ zu } B) \iff B = S^t AS \text{ f\u00fcr ein } S \in GL_n(\mathbb{R}).$$

Aus dem Spektralsatz folgt

$$A \sim B \iff \exists S \in O(n) : B = S^{-1}AS.$$

Damit ist die Ähnlichkeitrelation auf $S_n(\mathbb{R})$ eine feinere Äquivalenzrelation als die Kongruenzrelation.

Es gilt offensichtlich für $A, B \in S_n(\mathbb{R})$:

- $A \sim B \iff p_A = p_B$,
- $A \cong B \iff \operatorname{rg}A = \operatorname{rg}B$ und A und B haben dieselbe Anzahl positiver Eigenwerte (mit Vielfachheit gezählt).

Für “ \cong ” gibt es insbesondere nur endlich viele Äquivalenzklassen, nämlich $\binom{n+2}{2}$ Klassen. Für $n = 4$ kommen die 15 Diagonalmatrizen mit den Diagonaleinträgen $(0, 0, 0, 0)$, $(1, 0, 0, 0)$, $(-1, 0, 0, 0)$, $(1, 1, 0, 0)$, $(1, -1, 0, 0)$, $(-1, -1, 0, 0)$, $(1, 1, 1, 0)$, $(1, 1, -1, 0)$, $(1, -1, -1, 0)$, $(-1, -1, -1, 0)$, $(1, 1, 1, 1)$, $(1, 1, 1, -1)$, $(1, 1, -1, -1)$, $(1, -1, -1, -1)$, $(-1, -1, -1, -1)$ vor.

$(1, 1, 1, 1)$ liefert die vierdimensionale euklidische Geometrie. $(1, 1, 1, -1)$ ist der Minkowski-Raum, die Raumzeit der speziellen Relativitätstheorie.

Korollar 22.19 Ist (V, g) ein nicht-entarteter n -dimensionaler orthogonaler Raum über \mathbb{R} , (p, q) die Signatur von g , so ist die orthogonale Gruppe $O(V, g)$ isomorph zu der Matrixgruppe

$$O(p, q) := \{A \in GL_n(\mathbb{R}) \mid A^t E_{pq} A = E_{pq}\},$$

$$\text{wobei } E_{pq} := \left(\begin{array}{cccc} 1 & & & 0 \\ & \ddots & & \\ & & 1 & \\ & & & -1 \\ 0 & & & & \ddots & \\ & & & & & -1 \end{array} \right) \left. \begin{array}{l} \\ \\ \\ \\ \\ \end{array} \right\} \begin{array}{l} p - \text{Zeilen} \\ \\ \\ q - \text{Zeilen} \end{array}$$

Speziell ist $O(n) = O(n, 0)$ die orthogonale Gruppe des euklidischen Raumes.

Beweis: Es sei (e_1, \dots, e_n) eine Basis von V mit $g(e_i, e_j) = 0$ für $i \neq j$ und

$$g(e_i, e_i) = 1 \text{ für } 1 \leq i \leq p, \quad g(e_i, e_i) = -1 \text{ für } i = p+1, \dots, n$$

Ist $F : V \rightarrow V$ orthogonal (bzgl. g) so ist die Matrix A von F bezüglich der Basis (e_1, \dots, e_n) ein Element von $O(p, q)$, denn

$$\begin{aligned} F(e_j) = \sum_{i=1}^n a_{ij} e_i &\Rightarrow g_{ik} = g(e_j, e_k) = g(F(e_j), F(e_k)) \\ &= g\left(\sum_{i=1}^n a_{ij} e_i, \sum_{l=1}^n a_{lk} e_l\right) = \sum_{i,l=1}^n a_{ij} g_{il} a_{lk}. \end{aligned}$$

□

Die Gruppe $O(3, 1)$ heißt die **Lorentzgruppe**.

Übungen

1. Es sei (V, g) ein orthogonaler Raum über K . G sei die Matrix von g bezüglich einer Basis \mathcal{B} von V . Ist $\det G$ unabhängig von der Wahl der Basis?

2. Es sei (V, g) ein orthogonaler Raum über K . $r = rg(g)$. Welche Dimension hat V^\perp ? Was bedeutet $V^\perp = 0$?

3. Es sei $\cosh(x) = \frac{1}{2}(e^x + e^{-x})$, $\sinh(x) = \frac{1}{2}(e^x - e^{-x})$. Weiter sei

$$A_x := \begin{pmatrix} \cosh x & \sinh x \\ \sinh x & \cosh x \end{pmatrix}.$$

Zeigen Sie: $x \mapsto A_x$ ist ein injektiver Gruppenhomomorphismus

$$\Phi : (\mathbb{R}, +) \rightarrow O(1, 1)$$

und das Bild von Φ besteht aus allen Matrizen $\begin{pmatrix} a & b \\ c & d \end{pmatrix} \in O(1, 1)$ mit

$\det \begin{pmatrix} a & b \\ c & d \end{pmatrix} = 1$ und $d > 0$. Diese Gruppe wird mit $SO^+(1, 1)$ bezeichnet

Zeigen Sie $SO^+(1, 1) \cup \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} SO^+(1, 1) \cup \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} SO^+(1, 1) \cup$

$\begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} SO^+(1, 1) = O(1, 1)$ und dies ist eine disjunktive Vereinigung.

4. Es sei (V, g) ein nicht-entarteter n -dimensionaler orthogonaler Raum.

$\mathcal{B} = (v_1, \dots, v_n)$ eine Basis von V derart, dass die Untervektorräume $V_i = \text{Span}(v_1, \dots, v_n)$ nicht-entartet sind.

Dann kann man folgendermaßen eine Orthogonalbasis (w_1, \dots, w_n) von V finden, für die $\text{Span}(w_1, \dots, w_n) = V_i$ gilt:

Man setzt: $w_1 = v_1$ und für $i \geq 2$ sei $(x_1^{(i)}, \dots, x_{i-1}^{(i)}) \in K^{i-1}$ eine Lösung des linearen Gleichungssystems

$$\sum_{j=1}^{i-1} x_j^{(i)} g_{jk} = g_{ik} \text{ für } k = 1, \dots, i-1,$$

wobei $g_{ij} = g(v_i, v_j)$. Man setze dann

$$w_i = v_i - \sum_{j=1}^{i-1} x_j^{(i)} v_j.$$

Beweisen Sie, dass (w_1, \dots, w_n) eine Orthogonalbasis von V ist und

$\text{Span}(w_1, \dots, w_i) = V_i$ für $i = 1, \dots, n$ gilt.

Führen Sie diese Verfahren am Beispiel

$$(a) \quad (\mathbb{Q}^3, g), \quad G = \begin{pmatrix} 1 & -2 & 0 \\ -2 & 1 & \frac{3}{2} \\ 0 & \frac{3}{2} & 0 \end{pmatrix}, \quad g(x, y) = x^t G y,$$

$$v_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad v_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \quad v_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}.$$

$$(b) \quad (\mathbb{F}_5^3, g), \quad G = \begin{pmatrix} 2 & 3 & 4 \\ 3 & 4 & 2 \\ 4 & 2 & 3 \end{pmatrix}, \quad g(x, y) = x^t G y,$$

$$v_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad v_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \quad v_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}.$$

durch.

Zeigen Sie im Fall (b), dass G zu der Matrix $\begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ über \mathbb{F}_5 kongruent ist.

5. Es sei (V, g) ein n -dimensionaler nicht-entarteter orthogonaler Raum über K , und es sei (v_1, \dots, v_n) eine Basis, so dass $V_k = \text{Span}(v_1, \dots, v_k)$ für $k = 1, \dots, n$ nicht-entartet ist. Es sei $G = (g_{ij})$ mit $g_{ij} = g(v_i, v_j)$. Die Determinanten $\det G_k$ sind dann von Null verschieden, wobei $G_k = (g_{ij})_{i,j \leq k}$. Dann ist G kongruent zu der Diagonalmatrix

$$\begin{pmatrix} \det G_1 & & & 0 \\ & \frac{\det G_2}{\det G_1} & & \\ & & \ddots & \\ 0 & & & \frac{\det G_n}{\det G_{n-1}} \end{pmatrix}$$

Wie kann man im Fall $K = \mathbb{R}$ an den Determinanten $\det G_k$, $k = 1, \dots, n$ die Signatur von g ablesen?

Hinweis: Ist S obere Dreiecksmatrix und $S^t G S = \begin{pmatrix} a_1 & & \\ & \ddots & \\ & & a_n \end{pmatrix}$, so ist

$$(\det S_k)^2 \det G_k = a_1 \cdot \dots \cdot a_k.$$

Literatur

- [1] Artin, Emil: Geometric Algebra. Interscience Publishers 1957, Neudruck: Wiley Classics Library Edition, 1988.
- [2] Artin, E.: Analytische Geometrie und Algebra I,II (SS 1960, WS 1960/61, Hamburg).
- [3] Michael Artin: Algebra, Birkhäuser Advanced Texts, 1993
- [4] Brieskorn, Egbert: Lineare Algebra und Analytische Geometrie I und II, Vieweg 1983.
- [5] Curtis, Charles W.: Linear Algebra, an introductory approach. Allyn and Bacon, Inc., Boston, 1977, third printing.
- [6] Dixon, Crist: Linear Algebra. Van Nostrand Reinhold Company, New York, 1971.
- [7] Eisenreich, Günther: Lineare Algebra und analytische Geometrie. Akademie-Verlag, Berlin, 1980.
- [8] Fichtenholz: Differential- und Integralrechnung I, Harri Deutsch
- [9] Fischer, Gerd: Lineare Algebra. Vieweg Grundkurs Mathematik, Braunschweig, 1979, 5. Auflage.
- [10] Forster: Analysis I, Vieweg Grundkurs Mathematik
- [11] Grauert und Lieb: Differential- und Integralrechnung III (2. Auflage)
- [12] Grauert, H.: Analytische Geometrie und lineare Algebra I,II (WS 1972/73, SS 1973, Mathematisches Institut der Universität Göttingen).
- [13] Gröbner, Wolfgang: Matrizenrechnung. B.I. Wissenschaftsverlag, Mannheim, 1966.
- [14] Kolman, Bernard: Introductory Linear Algebra with Applications. Third edition, Macmillan Publishing Company, New York, 1984.
- [15] Kostrikin, A.I.: Introduction to Algebra. Springer-Verlag, 1982.
- [16] Kostrikin, A.I.; Manin, Yu.I.: Linear Algebra and Geometry. Gordon & Breach, 1989.
- [17] Kowalsky, Hans-Joachim: Lineare Algebra. De Gruyter + Co, Berlin, 1969.
- [18] Lang, Serge: Introduction to Linear Algebra. Addison Wesley, 1970.
- [19] Lang, Serge: Algebra. Addison-Wesley, Fourth Printing, 1971.
- [20] Lang, Serge: Linear Algebra. 3rd corr. printing 1993 (Undergraduate Texts in Mathematics), Springer-Verlag.

-
- [21] Smith, Larry: Linear Algebra. Springer-Verlag, New York, 1978.
 - [22] Sperner, Emanuel: Einführung in die Analytische Geometrie und Algebra (1. und 2. Teil). Vandenhoeck und Ruprecht Göttingen, 1948; 6. Auflage, 1963.
 - [23] Spindler, H.: Elementare Zahlentheorie, Vorlesungsausarbeitung SS 97
 - [24] Scheja, G.; Storch, U.: Lehrbuch der Algebra Teil 1 und Teil 2. Teubner, Stuttgart, 1988.
 - [25] Jürgen Schmidt: Mengenlehre I, B.I. Mannheim, 1966
 - [26] Trapp, H.-W.: Einführung in die Algebra (Osnabrücker Studien zur Mathematik).
 - [27] van der Waerden, B.L.: Algebra I und II. Springer-Verlag, Berlin und Heidelberg, 1967 (5. Auflage).

Index

- \mathbb{F}_p , 51
 \mathcal{E} , 218
 äquivalente Matrizen, 118

 nicht-entartet, 225

 Abbildung, 27
 abelsch, 192
 abzählbar, 33
 Addition, 46, 60, 140
 additive Gruppe, 193
 additive Inverse, 47
 adjungierte Operator, 215
 affiner Unterraum, 129
 Algebra, 136
 Algebromorphismus, 138
 algebraisch, 145
 algebraisch abgeschlossen, 169
 algebraische Gleichung, 4
 algebraischer Abschluß, 169
 algebraisches Komplement, 156
 Algebrasomorphismus, 138
 allgemeine Lage, 131
 alternierende Funktion, 150
 antilinear, 205
 Äquivalenzklasse, 42
 Äquivalenzrelation, 41, 42
 äquivalent, 16
 assoziativ, 46
 aufgespannter Untervektorraum, 63
 Auswertungsabbildung, 144

 Basis, 67
 Betrag, 56
 Bidualraum, 93
 bijektiv, 29
 Bild, 27, 83
 Bild von f , 28
 Bildmenge, 28
 Bilinearform, 174
 binäre Relation, 41
 Binomialkoeffizient, 38

 Cauchy-Schwarzsche Ungleichung, 178, 206
 Charakteristik, 50

 charakteristisches Polynom, 167
 Corang, 83
 Cramersche Regel, 157

 Definitionsbereich, 27
 Determinante, 13, 149, 167
 diagonalisierbar, 164
 diagonalisierbar über K , 164
 Differenzmenge, 26
 Dimension, 68
 Dirac-Matrizen, 109
 direkt, 65
 direkte Summe, 64
 disjunkt, 25
 Diskriminante, 178
 Distributivgesetz, 47
 Division mit Rest, 144
 Drehungen, 196
 Dreiecksgeometrie, 180
 Dreiecksungleichung, 179
 dritte Einheitswurzel, 5
 duale Abbildung, 88
 duale Basis, 89
 duale Paarung, 90
 Dualraum, 87
 Durchschnitt, 25

 Eigenfunktion, 221
 Eigenraum, 165
 Eigenvektor, 165
 Eigenwert, 165
 eindeutig lösbar, 15
 Einheit, 155
 Einheiten, 194
 Eins, 47
 Einschränkung, 30
 Einselement, 47
 elementare Umformungen, 16
 Elementarmatrix, 116
 Elemente, 23
 endlich, 32
 endlich erzeugt, 63
 endlichdimensional, 68
 Endomorphismenalgebra, 136

- Endomorphismenring, 136
Endomorphismus, 136
Entwicklung nach der k -ten Spalte, 158
Erwartungswert, 176
erweiterte Matrix, 15
erweiterter euklidischer Algorithmus, 53
Erweiterungskörper, 49
Erzeugendensystem, 63
euklidische Abstand, 190
euklidische Norm, 178
euklidischer Vektorraum, 174
exakte Sequenz, 90
- Faser von f über y , 28
Fermatsche Zahl, 37
Fixpunkt, 190
Fortsetzung, 30
Fourier-Polynome, 174
Fourierkoeffizient, 90, 183
freie Variable, 20
Fundamentalsatz der Algebra, 169
funktorielle Eigenschaften, 88
- ganze Zahlen, 23
Gaußsches Eliminationsverfahren, 19
geordnete n -Tupel, 27
geordnete Paar, 26
Grad eines Polynoms, 142
Graf, 27
Gramsche Determinante, 176
Gramsche Matrix, 176, 191
Graph, 41
Gruppe, 192
Gruppenhomomorphismus, 193
Gruppenisomorphismus, 193
- Hauptminoren, 187
Hauptuntermatrix, 187
Hauptvariable, 20
hermitesche Form, 205
homogen, 14
homogene Gleichungssystem, 125
Homomorphismus, 79
Hyperebene, 129
- identische Abbildung, 30
imaginäre Achse, 55
Imaginärteil, 55
- Induktionsanfang, 36
Induktionsschluß, 37
inhomogen, 14
inhomogenes Gleichungssystem, 125
injektiv, 29
Integritätsbereich, 134
invariant, 161
Invarianten, 168
Inverse, 47, 155, 192
inverse Matrix, 112
invertierbar, 112, 155
Isometrie, 190
Isometrien, 224
isomorph, 81, 193
Isomorphismus, 81, 224
isotrop, 225
- K -Algebra, 136
 K -Basis, 67
 K -Unteralgebra, 136
 k -fache Nullstelle, 170
Körper der Charakteristik p , 50
Körper der Charakteristik Null, 50
Körperautomorphismus, 53
Körpertheorie, 46
kanonische Abbildung, 138
kanonische Projektion, 43
kartesisches Produkt, 27
Kern, 83
Kodimension, 129
Koeffizienten, 14
Koeffizientenfolge, 140
Kofaktor, 156
kommutativ, 47
kommutativer Ring, 134
komplementäre Matrix, 156
komplexe Ebene, 55
komplexe Fourierpolynome, 206
komplexer Vektorraum, 60
Komponente, 15
Komposition, 30
kongruent, 230
kongruent modulo G , 144
Kongruenzabbildung, 190
konjugiert, 213
konjugiert komplexe Zahl, 56
konstante Polynome, 141

- konstante Term, 142
konstante Terme, 14
Kosinussatz, 180
Kroneckersymbol, 89
- lösbar, 15
Lösung, 15, 125
leere Menge, 24
Leitkoeffizient, 142
linear, 79
linear abhängig, 67
linear unabhängig, 67
Linearformen, 87
Linearkombination, 62
Lorentzgruppe, 231
- Markov-Matrix, 109
Matrix, 14, 98
Matrix einer linearen Abbildung, 103
Matrizenalgebra n -ten Grades, 136
Matrizenring n -ten Grades, 135
maximaler Rang, 112
Menge, 23
Minimumsprinzip, 36
Minkowski-Raum, 225
Monom, 142
multilineare Funktion, 150
Multiplikation, 46, 140
multiplikative Gruppe, 193
- natürliche Abbildung, 43
natürliche Zahlen, 23
natürlicher Isomorphismus, 94
negative Element, 47
negativen orthogonalen Raum, 225
nicht lösbar, 15
nicht-entartet, 224
nilpotent, 138, 146
Norm, 178
normiert, 170
Null, 47, 60
Nullelement, 47
Nullstelle, 145
Nullteiler, 137
nullteilerfrei, 48
nullteilerfreier Ring, 134
Nullvektor, 60
- Ordnung einer Nullstelle, 170
orthogonal, 181, 190, 223
Orthogonalbasis, 227
orthogonale Gruppe, 195, 224
orthogonale Komplement, 181
orthogonale Projektion, 182
orthogonale Projektor, 183
orthogonaler Raum, 223
Orthogonalraum, 223
orthonormal, 181
Orthonormalbasis, 181, 183
Orthonormalisierungsverfahren, 184
Orthonormalsystem, 183
- Paar, 26
Pauli-Matrizen, 108
Polynom, 139
Polynomfunktion, 71
positiv definit, 174, 205
positive orthogonalen Raum, 225
Potenz, 193
primitiv, 59
Produktmatrix, 106
Pythagoras, 180
- quadratische Ergänzung, 4
Quotientenbildung, 41
Quotientenmenge, 43
- Rang, 83, 112
rationale Zahlen, 23
Realteil, 55
reduzierte Zeilenstufenform, 127
reelle Achse, 55
reelle orthogonale Gruppe, 195
reelle Zahlen, 23
reeller Vektorraum, 60
Relation, 41
Repräsentant, 43
Rest einer Polynomdivision, 144
Reziproke, 47
Ring, 134
Ringhomomorphismus, 138
Ringisomorphismus, 138
- Satz von Thales, 180
selbstadjungiert, 215
Sequenz, 90

- Sesquilinearform, 204
Signatur, 229
skalare Multiplikation, 60
Skalarprodukt, 174
Spalte, 14
Spalte einer Matrix, 98
Spaltenrang, 111
Spektralzerlegung, 220
spezielle lineare Gruppe, 194
spezielle unitare Gruppe, 209
Spiegelungen, 196
Spingruppe, 210
Spur, 159, 167
Standardbasis, 71
Standardbilinearform, 102
Standardskalarprodukt, 102
Stufenform, 18
surjektiv, 29
symmetrisch, 174
symmetrische Gruppe, 194
- Teilmenge, 23
Torus, 45
Transformationsformel für Bilinearformen,
177
Translation, 190
transponierte Matrix, 104
trigonometrische Polynome, 174
Tupel, 27
- überabzählbar, 33
Umkehrabbildung, 29
unabhängig, 189
unendlich, 32
unitärer Vektorraum, 206
unitäres Skalarprodukt, 206
unitar, 209
unitare Gruppe, 209
Unteralgebra, 136
Untergruppe, 193
Unterkörper, 49
Untermatrix, 113
Unterring, 135
Untervektorraum, 61
Urbildmenge, 28
- Vektoren, 60
Vektorraum, 60
Vektorraumhomomorphismus, 79
Vereinigung, 24
volles Repräsentantensystem, 43
vollständige Induktion, 36
- Wahrscheinlichkeit, 176
Wahrscheinlichkeitsvektoren, 109
Wertebereich, 27
Winkel, 179
- Zeile, 14
Zeile einer Matrix, 98
Zeilenrang, 111
Zeilenstufenform, 18
Zufallsvariable, 176, 189